

THE METHOD OF TESTING LITERARY
APPRECIATION¹

BY CHARLES FOX

- I. *Introduction* (pp. 1-2).
- II. *The nature of the test* (pp. 2-3).
- III. *Testing the test* (pp. 3-6).
- IV. *Introspective examination* (pp. 6-9).
- APPENDIX I. *Samples of the literature tests* (pp. 9-10).
- APPENDIX II. *Table of correlations* (p. 11).
- References* (p. 11).

I. INTRODUCTION

THE factors involved in literary appreciation are so complex and interwoven that the first task of psychology is to make sure that any suggested test really does measure what it purports to measure. Various attempts have been made to devise tests of appreciation, and as will be easily realized the chief difficulty is to secure a valid criterion. Standards of literary worth have changed so much in recent times that one is almost forced to assert that there are no objective criteria. And certain writers have arisen who have tried to discard all the literary traditions of the classical writers. The attempt to measure literary taste is thus subject to a double difficulty; the selection of material that is considered of literary merit and the assessment of the individual's reaction to the material.

The most successful attempts to secure a test of literary appreciation have relied on the opinion of experts, such as literary critics, writers, university teachers and so on⁽¹⁾. These are asked to decide on the merit of selections of prose or verse; and the person tested is judged by his agreement with the experts, whose consensus of opinion is regarded as a sufficient objective criterion. The psychologist selects his material from classical sources or from sources generally approved and submits these to the experts. He selects two or more extracts of a kindred kind, such as descriptions of scenery. The experts range them in order of merit and

¹ A paper read before the XIth International Congress of Psychology at Paris, 29 July 1937.

the person tested has to put them in the same order to be considered as having literary appreciation (2). In one sense this method is entirely subjective in relying upon opinion; but as a test is only employed if the expert opinions agree in choosing it, this may be considered as a sufficiently objective criterion if there are adequate instructed opinions. Such, in fact, is always the method adopted in science for securing objective validity.

The earliest attempts at measuring appreciation proceeded on the assumption that specific tests ought to be used to analyse the various factors in appreciation. And the earliest material for dealing with appreciation was of the simplest kind, such as simple line diagrams, colours, tones, chords, simple rhythms and so on (3). It was hoped in this way to analyse appreciation into its factors and by assigning marks to each factor to get a total measure of the person's ability. Such a method repeats the mistakes made in the early attempts to measure intelligence. Here, too, it was hoped to get hold of some simple sensory factors such as sensory acuity by which intelligence could be evaluated. The attempt proved to be a failure; and Binet was the first to realize that to measure intelligence we must deal with complex processes in their undivided complexity, and that the material should be drawn from everyday experience instead of being artificially manufactured in the laboratory. The success of his method is the justification of his procedure. The analysis of intelligence is only just beginning and is the result of finding suitable tests of intelligence, not vice versa; that is to say the analysis comes after the testing has been found to be possible.

II. THE NATURE OF THE TEST

Similarly, I think that the psychologist's attempt to measure literary appreciation must begin with complex literary material and leave the analysis of the process until the tests themselves have been proved valid. I am not at all sure that the psychologist can ever succeed in the task of analysing appreciation unless he is himself a literary person. Nevertheless the simpler task of devising a test should not be beyond his power. For this purpose we must select our material from acknowledged classical sources of literature. It is not the psychologist's business to attempt to evaluate the worth of such literature. He must accept it until such time as it is overthrown by something which is equally widely acknowledged and has stood the test of time. He must assume that what Shakespeare, for example, wrote is the best literature, and that the judgement of time is for him final. If he is willing to make the further assumption that

writers of acknowledged reputation are the best judges of their own work his task is possible. By that I mean that if a great writer, in the sense I have used the term, employs a word or a phrase in a given context, the psychologist cannot except by chance make a better choice than the writer did.

With these assumptions in mind I have attempted to construct a test of literary appreciation using the devices which have proved successful in mental testing. My objective criterion is what the author wrote; and the subject tested has to try to select this from a number of alternatives. It may happen that my alternatives would now be considered better than what the author wrote but I feel bound to rely on his acknowledged reputation. Accordingly a number of short extracts from acknowledged authorities in verse and prose were selected, eight in verse and nine in prose. Each extract was complete in itself, needing no further context. The selections were printed with two words or phrases omitted from each, and below the selection the omitted word or phrase was printed with three other alternatives. The subjects were asked to underline the word or phrase most appropriate for the context. In a preliminary trial all those extracts were rejected which failed to secure more than 'chance' marks and all those which secured full or nearly full marks. The former were too difficult and the latter too easy. This left the 17 extracts, for each of which 2 marks could be obtained, making a maximum of 34 marks. Samples of the extracts are given in Appendix I.

In testing appreciation it is essential that the subjects should feel at their ease and not be pressed. Accordingly they were given as much time as they needed. No subject required more than 1 hr., and several completed the test in 20 min. It is advisable to make the suggestion that their judgement will be sounder if they do not take an unduly long time to make up their minds; but it should be emphasized that this is a suggestion merely and not an instruction.

III. TESTING THE TEST

The test was given to seventy-five men students all of whom had graduated in honours. They also had two tests of intelligence, Otis advanced examination and Cattell's test scale III, Form A. The subjects in which the students had graduated were recorded, together with the class they had gained in their degree.

The first step in the enquiry was to determine whether the test was actually a test of literary appreciation. It is too often assumed in such

investigations that the mere fact that the test deals with literary material is a sufficient guarantee that it is measuring literary attainment. But the ability to discriminate the correct word or phrase might be due mainly to general intelligence. The data that we have collected enable us to determine this point. It is a reasonable assumption that, if excellence in the tests is an indication of literary taste, those who have been trained in the University in literary subjects should do better than those whose training has not been literary. Our investigation has enabled us to test this assumption; and also the further assumption that the better the degree obtained by the literary person the better his literary taste. Now success in choosing one of four alternatives to fit into a given context may be primarily due, as was said, to an act of intelligence and have little to do with literary appreciation. The statistical sorting of the results has enabled us to answer this objection.

Among the group of seventy-five honours graduates there were fifty-five who had included literary subjects in their degrees, as they had read either English or Classics or Modern Languages or History; and there were twenty who had taken Science or Mathematics or some other non-literary subject.

The average mark obtained by the literary group in the literary test was 15.83 and by the non-literary group 14.35, yielding a difference of 1.48 which statistically is insignificant (as the s.d. of the difference is 0.945).

For the Cattell test of intelligence the average marks for the literary and non-literary men were 107.7 and 114.55 respectively with a significant difference in favour of the non-literary men of 6.85 (s.d. 2.737). Whereas for the Otis test the literary men had an average of 190.2 and the non-literary 201.95 with a significant difference of 11.75 (s.d. 3.906). The ratio of the standard deviation to the mean for the literary tests for the whole group was 0.234; whilst for the Cattell and Otis tests respectively it was 0.096 and 0.077. Hence these literary tests are considerably more discriminative in spreading the men than either of the mental tests. This result would indicate that literary taste is much more variable than intelligence and is an indirect proof that the tests are not merely measures of intelligence. An indirect confirmation of this is also obtained by grading the men according to the class they obtained in their degrees, giving 4 marks for first class, 2 for a second class first division, and progressively decreasing the marks until a third class man obtains $\frac{1}{2}$. The correlation coefficient between the marks so obtained and the marks in the literary test was 0.198 which is insignificant; so that if we assume

that, on the whole, the higher a man's intelligence the better degree he will obtain, the above result is confirmed. (These and other correlation coefficients are shown in Appendix II.)

Having assured ourselves that the tests are not simply new tests for measuring intelligence, have we any reason to believe that they measure literary appreciation? We may answer this question by separating the literary group from the others and assigning arbitrary marks to those parts of a man's degree subjects which are literary, and zero to non-literary subjects. In Cambridge it is the custom for a student to take two parts for his degree, which may be both literary parts, or both non-literary, or one of each. If we assign 4 marks to each literary part and 0 to non-literary parts a man may receive 8, 4 or 0 marks for his degree subjects. This enables us to calculate a correlation coefficient for the tests and the kind of degree that a student has taken. The correlation coefficient for the nature of the degree and the literary test score was 0.243, which is statistically significant. If we combine this device with the above method of weighting for a higher degree the coefficient rises to 0.285; and if we assign weights only to the excellence of the literary degrees the coefficient becomes 0.290. All these three coefficients are significant and suggest that the tests discriminate between those who have had literary training and those whose University course has been of a non-literary kind. It must be remembered that the non-literary group have all had some literary training at school, and many of them are interested in literature which they pursue as a hobby. No statistical estimate of this factor is possible but it is reasonable to suppose that if its effect could be allowed for the magnitude of the above correlations would be very considerably increased.

The next step is to see to what extent the results are affected by general intelligence. We may answer this question by the double weighting device last mentioned, using the method of partial correlations.

If we keep intelligence constant, the correlation between the literary tests and the effects of literary training, as estimated by the excellence of a student's literary degree, is 0.348 if we use the Otis test, and 0.376 if we employ the Cattell test. These numbers are both significant and mean that the odds against the result being due to chance are less than one in a hundred. If we take into account the fact previously stated that all the students, without exception, have had some measure of literary training either at school or in the University, it is evident that the tests give a clear indication that it is possible to measure literary taste. By adopting methods similar to those described in this paper it would be

feasible to devise a test suitable for school children by which a prediction could be made as to their possibility of benefiting by a literary training.

It is evident, in the light of our investigation, that any attempt to evaluate appreciation, whether literary, musical, artistic or any other kind must proceed on the principle here established. That is to say the tests must be shown capable of discriminating those who, on independent evidence, are already presumed to have the power. Otherwise the validity of the tests is unproven. Just as in intelligence testing the validity of the tests ultimately rests on discriminating those who are known by other independent criteria to be possessed of intelligence.

IV. INTROSPECTIVE EXAMINATION

In psychological investigations, especially where such purely subjective processes as appreciation are concerned, we are bound to take into account the results of introspection. Statistical data must be interpreted by concurrent introspective data. The subjects were not told that this was an investigation into literary appreciation. The intelligence tests had been given a week earlier and all they were told about these tests was printed at the head of them, as follows: "Literature Tests: In the following extracts the omitted word or phrase is placed below with three others. Underline the word or phrase that you consider most appropriate for the context. Read each extract twice before you make your choice." No further information was given though some of the subjects suspected that the test was one of appreciation. They were asked to state, if they could, what the tests measured in their own opinion. When the test was completed, and before it was marked, the students were asked to give as complete an introspective record as they could, but no hint was given as to what they were to look for.

An examination of these records reveals one fact which is of vital importance in such investigations, namely that practically all the subjects were keenly interested in the tests. Without such an attitude of mind an investigation of this kind is of little value.

The 'marks' ranged from 9 to 25, with a median of 14.2, upper quartile 17.1 and lower quartile 11.6. In the following analysis of some of the introspective records the letter L indicates a literary student, N a non-literary, and the number the mark he obtained.

L. 18. I was faced with two difficulties; first to find the sense of the passage; and secondly, to choose, if there was more than one phrase conveying the right

meaning, the most suitable words. I think I did this on two principles, by the feeling of the rhythm of the passages and by the sound of the words. In the poetry I think the latter was uppermost, but in the prose, the rhythm helped me more.... The style in which the passage was written sometimes helped to determine the appropriate phrase. For many of my choices I relied not so much on conscious thought as on what Americans call a 'hunch'; words felt right.... I found this a test of my aesthetic feelings, besides, and rather than, a test of my reasoning powers.

L. 14. My first feeling was that to one who has been trained in English for his degree the test is one of identification by style. To such this test would be no test of aesthetic appreciation but one of sheer knowledge of what words certain ages and certain poets would be likely to use.... It is a question of one man's choice against another. It may develop even into a realization of what best fits the context and tone of the *writer*, not what is ideally best.

L. 13. Pleasure caused by the examples.... Some appreciative ability required to find the most suitable word—a certain amount of reasoning required—test should reveal ability to appreciate.

$\frac{1}{2}$ N. $\frac{1}{2}$ L. 14. I was glad to have a literary test.... what was dominant in my mind throughout was the debate whether to follow my own aesthetic impulse or the word demanded apparently by the test.

L. 18. I enjoyed the test very much and found pleasure in seeing whether the phrase I chose fitted for its aptness of sense, its rhythm, its conformity with the style of the passage and so on.... As a test it seemed one of comprehension rather than appreciation; to realize that a passage was good or bad did not help one to fill in the blanks. What was necessary above all, perhaps, was a sense of rhythm.

L. 17. Primarily a considerable interest in the test.... I found difficulty in some cases in getting the right word. The author's conception might differ from one's own and it is not inconceivable that the reader might supply in some cases a better phrase than the author. In some cases none of the alternatives seemed satisfactory. In others two would seem equally fitting—one, perhaps, pleasanter in sound, the other a trifle apter in meaning. Some doubt in these circumstances as what exactly the test was testing.... There were some alternatives that emphatically should *not* go down; the putting down of these would be fair evidence of bad taste.

N. 14. Logic was found useful in some cases in ruling out some alternatives.... The selections were made, however, mostly because they 'sounded right'.... As a test I considered that in so far as such an intangible quality of the human mind as appreciation can be tested it was satisfactory.

L. 23. Although an English student I could not 'place' any of the extracts—except the prose passage from Dr Johnson.... In certain of the poetry extracts I was non-plussed as the choice was so difficult.... This is the most interesting test to date. It gives scope to the flair for picking out the right word or phrase. Reasoning enters into the test, as many of the choices can be eliminated by the exercise of reason, but room is left for the feel of the passage.

L. 22. The passages did not all seem to test the same faculty. In some one had to find the phrase which fits in best with the logical meaning, in others what is tested is the feeling for the sound, and the emotional association of words.... I found my

general attitude towards the test was a more respectful one than in previous tests [intelligence tests] partly because they seemed to call for a more difficult and vigorous mental effort.

L. 18. I found these tests extraordinarily fascinating to do. Yet despite this all the passages were difficult and in none of them did I feel satisfied that I had put down the right thing. I felt that more was demanded than the simple solution of the phrase carrying the right meaning. It seemed necessary to soak in the spirit of the passage and to imagine what words the writer would have chosen, judging by his style. . . . It seemed a test of appreciation of style. I enjoyed myself immensely, but feel very doubtful of the result.

L. 20. In general I found the test one of appreciation of the meaning of words and phrases (appropriate imagery). But logical deductions came into many of the examples; and, though *taste* was tested in some, on the whole I conclude that deduction from evidence is the chief factor tested.

L. 19. I found considerable difficulty in making up my mind, and a choice was only possible after a process of weighing and reasoning [he was one of the slowest]. These tests lay too much emphasis on the rational and intellectual faculties and cannot be considered as proper tests of the music and rhythm of poetry. . . . It is true that they are a good test of estimating aptness of phrase and accuracy of expression to suit the context, but I still feel that the process necessitated is far too logical and analytic.

N. 13. In approaching the test I felt that not being well versed in literature I should have difficulty, and this proved true. I attempted to work by a process of elimination but found that the tests were not workable in this way. The poems were fairly easy, but the prose proved exasperating. The general impression was that while my answers were sensible and satisfied me, the shades of difference between the other possibilities were so fine that it was quite impossible to be sure in the majority of cases. The questions were a test of appreciation of literature, form and substance—or that is what I felt they were supposed to be; actually they were in many cases a test of ability to guess correctly.

N. 13. Presumably the test is an attempt to discover one's power of appreciating literature. But surely the main result of the test will be to discover whether a person has received a critical training or not.

L. 19. This test held my attention far more than the previous intelligence tests, and I found it less fatiguing. All this test seems to discover are my personal preferences in literary matters.

L. 15. I found them primarily a test of perception of the emotional value of the words or phrases, i.e. which of the alternative phrases was most in harmony with the emotional tone of the passage.

L. 16. Verse: I cannot see what the tests are testing; because most of the alternatives *fit* the context—the selection of one is less a matter of *judgement* than of *feeling*. Prose: These seemed rather more useful, because the alternatives offered varied considerably in style and rhythm, so that there was real point in preferring one to another; and thus there was call for genuine literary judgement or at least taste. The test was one, as I said above, of a *feeling* for literary *aptness*—balance of rhythm, sound. In fact, as the instructions said, of literary appropriateness.

It will thus be seen that, on the whole, the literary subjects felt that the tests did bring into play their literary feeling whilst the non-literary subjects tended to exercise their critical judgement. This conclusion could not have been inferred from the statistical treatment of the results and shows the necessity in this type of investigation of collecting subjective as well as objective data. Whether the difference is the result of native endowment or training remains indeterminate.

APPENDIX I. SAMPLES OF THE LITERATURE TESTS

In the following extracts the omitted word or phrase is placed below with three others. Underline the word or phrase that you consider most appropriate for the context.

Read each extract twice before you make your choice. [The correct choice is italicized; and the author is named, though this is not given in the actual test.]

VERSE

1. *Marvel*

He without noise still travelled to his end,
As (1) suns to meet the night descend,
The stars that for him fought, had only power
Left to determine now his (2) hour,
Which since they might not hinder, yet they cast
To choose it worthy of his glories past.

(1) flaming, waning, transient, *silent*.

(2) *fatal*, natal, final, last short.

2. *Shakespeare*

The birds chant melody on every bush;
The snake lies rolled in the cheerful sun;
The green leaves (1) with the cooling wind
And make a (2) on the ground.

(1) rustle, *quiver*, wither, shiver.

(2) feigned darkness, moving pattern, *chequered shadow*, elfin twilight.

3. *E. B. Browning*

With stammering lips and insufficient sound,
I strive and struggle to deliver right
(1) day and night
With dream and thought and feeling interwound,
And only answering all the senses round
With octaves of a mystic depth and height
Which step out grandly (2) from
From the dark edges of the sensual ground.

(1) my innermost emotion, the noblest aspirations, *the music of my nature*, whate'er betokens goodness.

(2) to observe the light, and attract the sight, well within the sight, *to the infinite*.

PROSE

1. *Scott*

I can very seldom think to purpose by lying perfectly idle, but when I take an idle book, or a walk, my mind strays back to its task out of contradiction as it were; the things I read become mingled with those I have been writing, and something is concocted. I cannot compare this (1) to anything save that of a woman to whom the mechanical operation of spinning serves as a (2) to the songs she sings, or the course of ideas she pursues.

(1) *process of the mind*, function of the soul, subconscious integration, happy circumstance.

(2) voiceless echo, undercurrent, *running bass*, counterbalance.

2. *Johnson*

If by a more noble and more adequate conception, that be considered as wit which is at once natural and new, that which, though not obvious, is upon its first production, (1) ; if it be that which he that never found it, wonders how he missed; to wit of this kind the metaphysical poets have seldom risen. Their thoughts are often new, but seldom natural; they are not obvious, but neither are they just; and the reader, far from wondering that he missed them, wonders more frequently (2).

(1) recognised as original; felt to be to the point; *acknowledged to be just*; recognised to be able.

(2) by what process they were found; how they could ever be discovered; why their stupidity did not make them rejected; *by what perverseness of industry they were ever found*.

3. *Kinglake*

For the rest of the day we saw no human being; we pushed on eagerly in the hope of coming up with the Bedouins before nightfall. Night came, and we still went on in our way till about ten o'clock. Then (1) and the weariness of our beasts (they had already done two good days' journey in one), forced us to determine upon coming to a standstill. Upon the heights to the eastward we saw lights; these shone from caves on the mountain side, inhabited, as the Nazarene told us, by rascals of a low sort—not real Bedouins—men whom we might frighten into harmlessness, but from whom (2).

(1) our tiredness and gloom; *the thorough darkness of the night*; our wretched state of mind; the cold and the damp.

(2) nothing good could come; there was no hope for mercy; all decent men would shrink; *there was no willing hospitality to be expected*.

APPENDIX II. TABLE OF CORRELATIONS¹

	<i>75 pairs</i>					
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>
<i>c</i>	0.755	0.800	—	—	—	—
<i>d</i>	0.198	0.243	0.285	—	—	—
<i>e</i>	0.191	-0.232	-0.038	0.381	—	—
<i>f</i>	0.244	-0.283	-0.040	0.297	0.700	—
<i>c</i> ¹	—	—	—	0.290	-0.143	-0.122

The figures in italics are non-significant.

a = marks for degree class.

b = marks for degree subjects.

c = *a* + *b*.

*c*¹ = *a* + *b* for literary men.

= 0 for non-literary men.

d = score in literary test.

e = score in Cattell test.

f = score in Otis test.

$r_{c^1d.f} = 0.348$, $r_{c^1d.e} = 0.376$.

¹ I have to thank Mr G. B. Hey, B.A., for his kindness in working out this table.

REFERENCES

- (1) CARROLL, H. A. (1932). "A standardization test of prose appreciation, etc." *J. Educ. Psychol.* xxiii, 401.
- (2) FOX, C. (1928). *Practical Psychology*, Ch. vii. London: Kegan Paul.
- (3) Edited by HAMLEY, H. R. *The Testing of Intelligence*, Ch. xii. London: Evans Bros.

THE ACQUISITION OF VERBAL REPETITION HABITS

BY R. C. OLDFIELD AND O. L. ZANGWILL
Cambridge Psychological Laboratory

- I. *Introduction* (pp. 12-15).
- II. *Method* (pp. 15-16).
- III. *Procedure* (pp. 16-18).
 - (1) *Material* (pp. 16-17).
 - (2) *Learning* (p. 17).
 - (3) *Subjects* (p. 18).
- IV. *Methods of analysis* (pp. 18-20).
- V. *Results* (pp. 20-22).
- VI. *Discussion* (pp. 23-25).
 - (1) *The normal material* (pp. 23-24).
 - (2) *The ungrammatical material* (pp. 24-25).
 - (3) *The inconsistent material* (p. 25).
- VII. *Summary and conclusions* (pp. 25-26).
References (p. 26).

I. INTRODUCTION

THE object of this study is an experimental investigation of the relationship between verbal repetition habits and normal recall. The problem has a certain theoretical significance in view of the sharp distinction which has sometimes been drawn between 'pure memory' on the one hand, and 'the operation of motor habit' on the other. Thus Bergson⁽²⁾ envisages 'habit' as a function exclusively dependent upon specific motor mechanisms, whereas ordinary recollection he regards as a more purely and specifically psychical activity. His view has been adopted in all essentials by McDougall⁽⁶⁾, who differs from Bergson only in the interpretation which he places upon the so-called 'pure' memory function.

Experimental data bearing directly upon this problem are somewhat meagre. Smith & McDougall⁽⁷⁾, it is true, have carried out experiments which appear to support Bergson's theoretical distinction. They claim that tests "of what Prof. Bergson calls 'pure memory'...show high correlation", whereas "those tests showing all the characteristics of

habit formation . . . correlate with each other, but not with either member of the other group". In an examination of the recall of a prose passage, Jones & English⁽⁵⁾ have observed that "there is a distinct discrepancy between the ability to memorize the sense of material and to memorize the exact words". This result they consider to be antagonistic to any theory of a common memory ability, and claim that the burden of proof is thereby put "upon those who deny the Bergson-McDougall hypothesis of a difference between 'memory' and the mechanics of habit formation by contiguity association". Heymans⁽⁴⁾, on the other hand, finds the correlation between accuracy in picture description and in the reproduction of nonsense syllables to be distinctly high, far higher, indeed, than that between descriptions of pictures and the recall of narrative prose. It is claimed that this result must inevitably reduce the plausibility of the Bergsonian theory.

More recent work by English *et al.*⁽³⁾ once again suggests that there may be a fundamental difference between verbatim reproduction and the recall of 'substance'. These workers attempted to compare the relative facility of 'verbatim' and 'substance' learning in respect to the same task. They find that the retention curves for 'verbatim' learning show a typical Ebbinghaus decline, whereas those for 'substance' learning in most cases manifest a distinct rise. In conclusion, they suggest that learning involves "the close and inseparable cooperation of at least two processes: the one emphasized in verbatim learning and in the conditioned response, the other in 'insight', the learning of substance" (op. cit. p. 259). This latter factor they consider to be "tied up with the organized or coherent or meaningful in experience" (op. cit. p. 258) and to differ both from the learning of nonsense syllables and from the so-called 'logical' learning in which there is acquisition of coherent narrative as a repetition habit.

Bartlett⁽¹⁾ has also hinted at a difference between verbal habit and everyday recall. He points out that the Ebbinghaus technique of rote learning followed by verbatim recall is apt to result in the study of the acquisition and maintenance of specific repetition habits rather than in elucidation of the processes involved in normal remembering. As is well known, Bartlett's own approach involved the utilization of meaningful and relatively familiar types of material, presented under conditions analogous to those obtaining in ordinary life. He attempted, moreover, to dispense with the explicit presence of any specific "intention to remember"—a factor which had played so large and unsuspected a role in many of the earlier researches. Bartlett's results demonstrate clearly

the extreme inadequacy of any strictly reproductive conception of memory, but appear to find satisfactory interpretation in terms of a 'schematic' theory of experimental organization adopted in conjunction with a reconstruction theory of remembering.

Bartlett does not attempt to elucidate in detail the relations between verbal habit and constructive remembering. He does, however, imply that reproductive recall is to be envisaged as dependent upon mechanisms fundamentally similar to those which he supposes to mediate high-grade constructive response. "Rote memory", he writes, "...is the most natural way of maintaining a completed 'schema' as far as possible undisturbed" (op. cit. p. 203).

It would appear to the present writers that the most significant feature in the entire controversy is the question of meaning, and its relation to memory in general. For McDougall, at all events, the formation and reproduction of a verbal habit is to be considered as depending pre-eminently upon the establishment and re-excitation of "links of mechanical association" (6, p. 336). On the other hand, everyday recollection, the 'pure memory' of Bergson, is to be attributed almost wholly to the potency of meaning—"a purely psychical product of psychic activity"¹—as an effective determinant of reproduction. Bartlett has, however, pointed out that it is impossible to rid stimuli of meaning so long as they remain capable of arousing any human response (1, p. 4); it follows that the nonsense material so generally employed in the earlier memory experiments must acquire meaning of a somewhat specialized character. Bartlett's trenchant criticisms, moreover, suggest not only that it would be unwise to base a rigid distinction between 'memory' and 'habit' upon this factor of meaning alone, but also that an investigation into the significance of meaning for the acquisition of verbal habits might prove extremely pertinent in regard to the general problem at issue. A secondary object of our investigation will therefore consist in a preliminary assessment of the significance of meaning for learning and recall.

We may point out that in connexion with material such as a continuous prose passage, 'meaning' has at least two functionally distinct aspects. In the first place, one might take the meaning of such a prose passage to be the situation, or set of

¹ Op. cit. p. 337. Compare also p. 341: "The essential difference between the remembering of these two kinds was that in the one case meaning was at a minimum, and remembering depended almost wholly upon mechanical or neural association of the nature of habit; whereas the complex scenes and events remembered (in some instances after a single perception) were full of meaning."

events, which it describes, and this will obviously include reference to the objects, persons and places mentioned, the associative significance which these elements possess for the reader, and so forth. We may call this the *extrinsic meaning* of the passage. In the second place, there is the formal structure of the passage, its grammatical, syntactical and logical coherence, without which it could possess no extrinsic meaning at all. This we may call its *intrinsic meaning*. It is the possession of intrinsic meaning that distinguishes a prose passage from a chaotic collection of words.

In view of these considerations, it is proposed to work not only with normal, meaningful, grammatical material, but also with material specially prepared so as to manifest abnormalities either in regard to extrinsic or to intrinsic meaning. In the first place, passages were included in which the meanings of words and sentences were grossly distorted through the introduction of errors and inconsistencies in grammar and syntax. In such cases, we may legitimately suppose that it is the intrinsic meaning which undergoes maximum disorganization. In the second place, a form of narrative was so constructed as to contain inconsistencies of logical and factual coherence. Thus, although the grammatical structure was left undisturbed, the events described were reduced to chaos. Apprehension and acquisition of material of this type might be expected to involve violation of habitual schemes of extrinsic meaning.

It is hoped that a comparison of the recalls obtained at various stages during the learning and forgetting of the specially constructed material may throw light upon the significance for memory of the meaning factor in both of its aspects distinguished above.

II. METHOD

The method of investigation was as follows: three short narrative prose passages (A, B and C) averaging about seventy words in length were used. Three versions of each passage were prepared on the following plan. One version was 'Normal' (N), another 'Ungrammatical' (U), and a third 'Inconsistent' (I). In the N version, the grammatical construction was correct, and the factual sequence of a straightforward nature and familiar type. In the U versions wanton distortions in the grammatical form were introduced, but no further liberties taken with the plot. These distortions consisted of reversals of singular and plural and genders, and of disagreement of parts of speech. In the I versions, while the grammar was left undisturbed, far-reaching inconsistencies of factual and logical coherence were introduced. The subjects were divided into three equal groups (I, II and III). Members of each group were required to learn by heart three passages in accordance with the following scheme:

Group I	Group II	Group III
AN	BN	CN
BU	CU	AU
CI	AI	BI

At various stages during the processes of learning and forgetting they

were asked to recall as much of the material as they were able. The occurrence of promptings, errors and omissions were carefully recorded, and subsequently analysed in detail. By this means it was hoped to obtain information respecting the manner in which these verbal habits are built up and decay. The relevance of such data to the problem under consideration will be discussed later. The present paper is concerned only with the process of acquisition. Disintegration will be dealt with in a later paper.

III. PROCEDURE

(1) *Material*

AN. Under a cloudless sky, he left Royston at half past seven in the morning, the lately risen summer sun beating upon his face and almost blinding him. Walking rapidly in a north-easterly direction, he reached Cambridge in time for an early lunch. By suppertime he was at Ely, and waiting only to enjoy the evening glow of gold upon the two great western towers, he went to bed.

BN. The cook was not feeling well that night and kept her own gas fire on all night, turned low admittedly, but burning continuously. Now she stated that she lighted that fire about half past nine, when she went to bed. She swears positively that she heard ten o'clock strike after she was in bed.

CN. The wall is very old—just loose stones—and as we have never repaired it, it is in some places just a pile of stones. He was three or four yards away, to my left, when I started to clamber over, carrying my gun under my right arm. I remember that he had taken his foot down and was moving back a little when suddenly a stone slipped right under my right foot,—and I lurched forward. To my horror one barrel of my gun went off.

AU. Under a cloudless skies, he leave Royston at halves past seven in a morning, the lately rose summer suns beat upon his faces, and almost blind him. Walk rapidly in a north-easterly directions, he reach Cambridge in times for a early lunches. By suppertimes he are at Ely, and waited only to enjoyed a evening glows of gold upon the two great western tower, he gone to bed.

BU. The cook were not felt well that night and keeps her own gases fire on all nights, turning low admittedly, but burned continuously. Now she state that she lighting this fires about half past nines, when she go to beds. She swore positively that she hear ten o'clock struck after she am in beds.

CU. The walls am very old—just loose stone—and as we has never repairing it, it are in some place just a piles of stones. He is three or four yard away to my left, when I starts to clambered over, carry my gun under my right arms. I remembers that he have takes his foots down, and are moved back a little when suddenly a stones slipping right under my right foots and I lurching forwards. To my horrors one barrels of my gun going off.

AI. Under a cloudy sky, he left Royston at half past one in the evening, the lately set winter sun beating down upon his face and almost blinding him. Running slowly in a south-easterly direction, he reached Cambridge too late for an early tea. By

lunch-time he was not at Ely, and not waiting only to enjoy the midnight glow of gold upon the three small eastern towers, he rose from bed.

BI. The cook was not feeling well that night and kept her own gas fire off all night, turned high admittedly, but burning continuously. Now she stated that she lighted that fire about half past eleven when she went to bed. She swears positively that she heard nine o'clock strike after she was in bed.

CI. The wall is very new—just loose stones—and as we have always repaired it, it is in every place just a pile of stones. He was three or four hundred yards away, to my left, when I stopped clambering under, carrying my gun under my right leg. I forget that he had taken his foot up and was moving back a little when gradually a stone slipped right under my right hand and I lurched backwards. To my joy three barrels of my gun went off.

(2) *Learning*

The learning of all three passages by each subject was, wherever possible, accomplished in one sitting. The order of learning was N first, followed by U and then I. A brief pause occupied with conversation was allowed between each passage and the next, and at such other times as seemed desirable. The subject was handed a typed copy of the passage to be learnt, and instructed to read it through carefully three times at his own normal speed with a view to getting it by heart, and to hand the paper back when he had finished. He was then asked to repeat orally as much of the passage as he was able, and told that he would be prompted and corrected wherever necessary. At the end of this attempt the paper was returned to him with the request to read it through again from one to three times according to the degree of proficiency shown in the repetition. This procedure was repeated until the number of errors and prompts had reached an extremely low level. One experimenter directed the activities of the subject, prompting and correcting where necessary, whilst the other recorded, in as great a detail as was possible using longhand, errors, prompts and times taken in learning and repetition. To ensure as far as possible uniformity of method in recording, the experimenters occupied the same roles throughout.

While the number of readings demanded before each repetition (except the first) was within narrow limits variable and adjusted to the rate of progress of each individual, it was decided that in view of such individual variability, proportional progress from one trial to the next would be approximately the same for all subjects. The actual number of times that the subject read through the material before each repetition was therefore, for present purposes, neglected.

(3) *Subjects*

Fifteen subjects learnt the material. These were arbitrarily divided into three groups of five. Among these were three professional actors well experienced in memorizing prose material, a lecturer and three research workers in psychology, an anthropologist, and a number of undergraduates. Of the fifteen subjects, four were women.

IV. METHODS OF ANALYSIS

In view of Bartlett's results, it was decided that some such classification of errors as the following would be best suited to the purpose in hand:

- (1) Prompts.
- (2) Omissions.
- (3) Synonyms and Equivalents.
- (4) Factual Conventionalizations.
- (5) Verbal Rationalizations.
- (6) Ungrammatical Constructions.
- (7) Constructive Importations, Additions and Verbal Transpositions.
- (8) Unclassified.

The complexity of the reproductive processes is such that these categories cannot be regarded as being strictly segregated one from another, and we believe that a rigid system of classification is apt to result in theoretical artificiality. The following broad definitions should serve adequately to illustrate this system of classification:

(1) *Prompts*. In cases where a subject apparently found himself totally unable to begin or continue with his repetition, one or two of the required words were spoken by the experimenter. If the subject remained unable to continue, further words were given until the subject was able to proceed unaided, each such interruption by the experimenter constituting a further prompt for the purposes of scoring.

(2) *Omissions*. The following types of error were classified as constituting a single omission: (a) Omission of single parts of speech not involving interruption of delivery or any considerable distortion of the original meaning. Instances of omission that caused a significant change of meaning were scored not only as omissions but also as rationalization or conventionalization as the case might be. (b) Groups of words constituting relatively unified meaningful points. Thus the omission of a

subordinate clause did not count according to the number of words it contained but according to the number of such meaningful points.

Examples. (a) "...when a stone slipped..." for "...when suddenly a stone slipped..." (*simple omission*).

(b) "...carrying my gun under my right arm..." entirely omitted (*scored as two omissions*).

(3) *Synonyms and Equivalents*. This category comprises not only synonyms in the strictly semantic sense, but also a number of verbal transformations which in that particular context could be regarded as subserving the subject's intention to formulate a correct apprehension of the original meaning. The text of the original and the subject's formulation could in such instances be regarded as psychologically equivalent.

Examples. "...walking swiftly..." for "...walking rapidly..." (*strictly semantic type*). "...a step backwards..." for "...back a little..." (*equivalent type*).

(4) *Factual Conventionalization*. Bartlett has defined *rationalization* as reduction of material to a form that can be readily and satisfyingly dealt with by the subject⁽¹⁾, pp. 93-4). Where such reduction involves translation into a form that is "socially acceptable" (and indeed "socially determined") the process has been termed *conventionalization*. In the present circumstances, such conventionalization is shown by the manner in which descriptions of events and situations regarded as unfamiliar or unusual by the subject are translated into a form more in keeping with his everyday experience and social setting. The following types of error were therefore classified as factual conventionalizations:

Examples. "...by lunch time he was at Ely..." for "...by lunch time he was *not* at Ely..." (*scored also as an omission*). "...one barrel of my gun went off..." for "...three barrels of my gun went off..."

(5) *Verbal Rationalization*.¹ This category comprises only certain types of error made in the repetition of the ungrammatical passages. Typically this involved simple grammatical correction. Such correction was generally confined to alteration of the inflexion of single words, but might on occasion manifest itself in widespread change of grammar and sequence.

Examples. "...a pile of stones..." for "...a piles of stones..." (*simple type*). "...under my feet..." for "...under my right foots..." (*complex type*).

¹ This term has been used in a somewhat different and more comprehensive sense by Bartlett⁽¹⁾, p. 93).

(6) *Ungrammatical Constructions.* In the repetition of the ungrammatical passages, subjects were apt to introduce grammatical errors at places where none had existed in the original material.

Example. "...he am..." for "...he is..."

(7) *Constructive Importations, Additions, and Verbal Transpositions.* This category included a number of closely related types of essentially constructive change.

Examples. "...at half past nine..." for "...about half past nine...". "...Now the cook was not feeling well..." for "...The cook was not feeling well...". "...heard the clock strike ten..." for "...heard ten o'clock strike..."

(8) *Unclassified.* A very small number of errors occurred which could not readily be placed in any of the above classes. These were minor changes resulting from the operation of the general constructive tendency to organize material into a form satisfactory to the subject.

V. RESULTS

Table I shows the actual number of each type of error made at each repetition for each of the three types of material. Table II shows the distribution of each type of error expressed as a percentage of the total errors made in each type of material. Table III gives for each type of material the total of changes other than those found only in the Ungrammatical Material, that is to say Omissions, Transpositions, etc., Synonyms, and Factual Conventionalizations. Fig. 1 shows the manner in which Prompts become eliminated throughout the learning process. Fig. 2 shows the elimination of Factual Conventionalizations, and Fig. 3 that of the total constructive changes. Fig. 4 demonstrates the elimination of Verbal Rationalizations in the case of the Ungrammatical Material.

It will be noticed from Table I that elimination of Omissions proceeds in an essentially similar way for all three types of material. In the case of Transpositions it can be seen that, as compared with the other two classes of material, a great many errors of this type occur in the case of the normal material at all except the final stages of the learning process. On the other hand, the elimination of the type of error which we have called 'synonymous' progresses in a somewhat irregular manner. At the first repetition, however, a greater number by about 40% of this type of error occurs in the inconsistent material than in the other two types. Some striking effects are noticeable by a study of the values obtained

for Factual Conventionalization at the various stages in the learning process (see Fig. 2). It must be borne in mind, however, that our results are insufficient to admit of conclusions of a statistically valid nature. Two points are of special interest in connexion with the elimination of conventionalization. First, the high initial level and very rapid fall of

Table I. *Number of classified errors made at each repetition for each type of material*

Category	Material	Trial										Total
		1	2	3	4	5	6	7	8	9	10	
Omissions	N	37	18	14	11	6	9	1	0	—	—	96
	U	37	32	16	16	7	4	2	2	2	1	119
	I	29	19	17	3	4	1	2	0	—	—	75
Transpositions, etc.	N	55	46	43	31	22	11	3	6	—	—	217
	U	34	37	20	15	7	8	5	3	1	1	131
	I	44	24	17	11	9	2	1	2	—	—	110
Synonyms	N	5	6	3	5	3	2	—	—	—	—	24
	U	5	4	5	1	2	2	1	1	1	0	22
	I	8	3	3	3	3	0	—	—	—	—	20
Conventionalization	N	10	7	7	4	5	0	0	0	1	—	34
	U	6	12	5	4	0	1	—	—	—	—	28
	I	19	11	4	3	4	0	—	—	—	—	41
Rationalization	U	78	57	40	30	20	15	8	10	6	4	268
Ungrammatical Construction	U	14	6	2	3	4	2	0	0	0	0	31

Table II. *Percentages of total errors falling under each category for each type of material*

Group	Omissions	Trans-positions	Synonyms	Convention-alization	Rationali-zation	Ungram-matical Construction
N	25.9	58.6	6.5	9.2	—	—
U	19.9	21.9	3.7	4.7	44.7	5.2
I	30.5	44.7	8.1	16.7	—	—

Table III. *Total changes for each type of material*

Material		
N	U	I
371	300	246

errors of this type in the inconsistent versions: secondly, the low initial level in the ungrammatical material followed by a 100 % increase at the second trial. It is interesting to note that in the total constructive changes, whereas an average fall of about 35 % is recorded in the cases of the normal and inconsistent material between the first and second trials, the corresponding value for the ungrammatical material remains almost constant, indeed it shows a slight rise (Fig. 3).

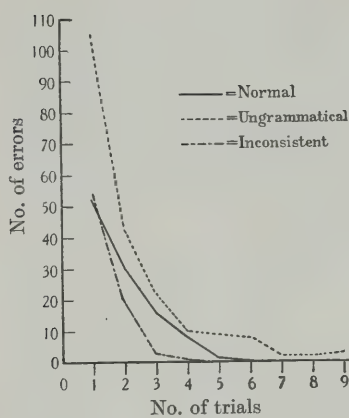


Fig. 1. Prompts.

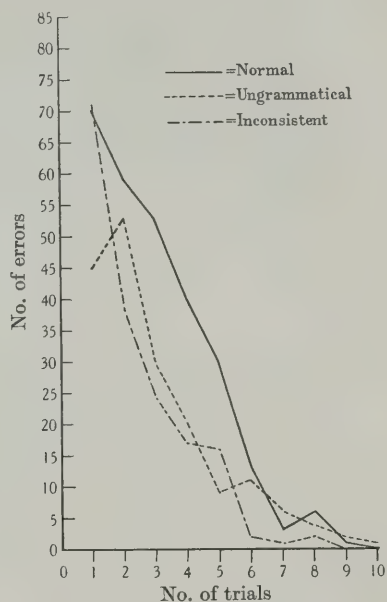


Fig. 3. Total "constructive" changes.

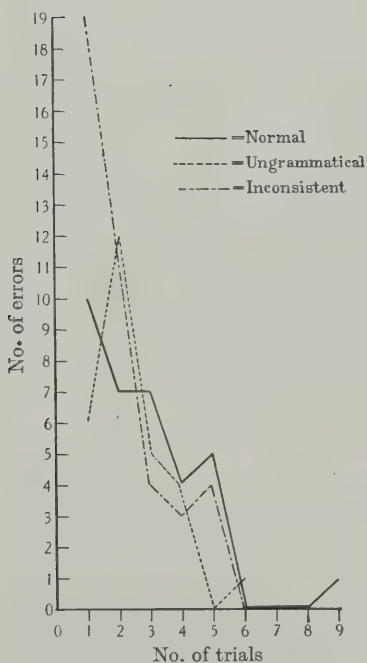


Fig. 2. Factual Conventionalization.

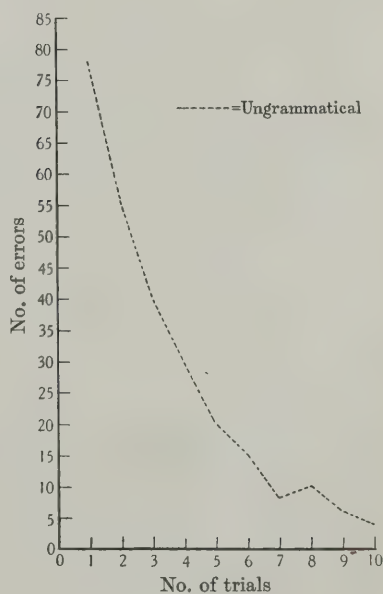


Fig. 4. Rationalization.

VI. DISCUSSION

(1) *The normal material*

The early repetitions possess an essentially constructive character. Many errors that occur will be seen to fall into categories that have been held to represent constructive types of change. As learning proceeds, errors of this type become progressively eliminated until a stage of almost literal reproduction is attained. From the nature of the curves, this transition has every appearance of being a continuous process. One may therefore suggest that with material of this type and under the conditions employed, the constructiveness of recall diminishes progressively during learning. The earliest repetitions indicate that a general scheme of the material, together with a few specific details, is all that is retained. Owing to our conditions of recall, these few details only can be accurately reproduced, but by the aid of prompts a continuous though as a rule extremely constructive version is rendered. One may suppose that with increase in the organization of the 'schema',¹ reproduction becomes more literal, and constructive tendencies operate to a much smaller extent.

From this account it may be concluded that verbal repetition habits are of a fundamentally 'schematic' nature. Interpretation on the basis of the earlier type of trace theory would require that improvement in performance should consist only in the progressive addition of detail at each stage, and in progressively greater availability of already acquired detail for the purposes of recall. A trace theory cannot account for the early formation of a general scheme. Our experiments have shown that the sequence of events involves at least the following stages: First, the formation of a very general 'scheme' with retention of 'dominant'² detail. At this stage, constructiveness operates in virtue of the existence of 'scheme' and detail, which, moreover, constitute the very material upon which constructive processes work. After a few more repetitions, it would appear that the 'scheme' becomes more articulated, comprising within itself more interrelated detail, and the essentially constructive character of remembering is consequently much less in evidence. At the final stages, a degree of organization of the scheme is attained sufficient

¹ Bartlett's definition of 'schema' as "an active organization of past reactions, or of past experiences, which must always be supposed to be operating in any well-adapted organic response" ((1), p. 201) will be provisionally accepted. A more detailed discussion of this concept will be presented in a forthcoming paper by the present writers.

² For explanation of the use of this term, see Bartlett((1), pp. 31-2).

to permit of almost literal reproduction. This may, so far as our evidence goes, be interpreted in either of two ways. In the first place one might regard this verbatim reproduction as representing the autonomous functioning of a well-organized 'scheme', constructive processes intervening only in the event of a breakdown of the reproductive mechanism. But, on the other hand, it might be supposed that the 'schematic' material available for constructive remembering is so well organized and differentiated that reconstruction can be *identified* with literal reproduction. Whichever of these alternative views may be correct, there can be no doubt that the acquisition of verbal repetition habits is a process involving the establishment, differentiation and operation of schematic mechanisms. This is in fact the view that Bartlett himself appears to adopt,¹ although, as has previously been pointed out, the actual relation between verbal repetition habits and normal remembering has not been worked out by him in detail.

(2) *The ungrammatical material*

From analysis of recall of the ungrammatical material the following facts have emerged: (1) Grammatical Rationalizations constitute a very high percentage of the total errors made. (2) The absolute number of errors falling under what we have termed constructive categories is lower by 40 % than in the normal passages. (3) During the earlier trials constructive errors fall off at a very much lower rate than in either the normal or the inconsistent material. In the case of Conventionalizations, in fact, there is a *rise* amounting to 100 % between the first and second trials. (4) 5 % of the total errors in this kind of material are of the type called Ungrammatical Constructions. (5) The number of Prompts is very much larger in the ungrammatical material than in either of the other two types.

These facts indicate that the process of acquisition of ungrammatical material as a repetition habit takes the following course: First, it is unlikely that any but the very vaguest scheme is in operation at the first repetition. This is suggested by the high value for Prompts. A few phrases, generally characterized by dominant detail, can be recalled, though often in a strongly rationalized fashion. Nevertheless, an awareness of the ungrammatical nature of the original material must form part

¹ See Bartlett(1), p. 205. Prof. Bartlett has, however, pointed out to us that a highly articulated and differentiated scheme might perhaps be identified with an organized system of traces. This possibility will be discussed by the present writers in a subsequent communication.

of this scheme, since Ungrammatical Constructions are very much in evidence at the first repetition. At the second repetition, a more coherent version of the original material is obtained; Prompts and Rationalizations diminish, and recall has that constructive character which typified the *first* repetition of the *normal* material. Differentiation and organization of the 'scheme' then takes place in a manner similar to that which we have already discussed. The lower level of constructive changes found in material of this type probably indicates that some at least of the subjects adopted a rather different method of learning this material as compared with normal. It would seem that intrinsic meaning is of considerable importance in the early stages of 'scheme' formation, and must be taken into account in any theory of remembering.

(3) *The inconsistent material*

The following are the principal facts relating to the recall of the inconsistent material: (1) The total of errors is lower than in either of the other two types of material. (2) The initial level both of Factual Conventionalizations and of synonymous change is higher than are the corresponding values for the other two types of material.

The first of these two results may well be due to the influence of practice in rote learning, since the order in which the passages were learnt was always the same, namely (1) Normal, (2) Ungrammatical, and (3) Inconsistent. As is to be expected, material incorporating factual and logical inconsistencies provides opportunity for widespread operation of the particular type of constructive change represented by conventionalization and equivalent synonyms. It can safely be concluded that the peculiar character of the material does not interfere with the process of scheme formation. Indeed certain bizarre elements may actually function as dominant and facilitate recall. It would seem that disturbance of the mode of extrinsic meaning, in a direction divergent from the habitual, exercises little effect upon the process of scheme formation.

VII. SUMMARY AND CONCLUSIONS

An attempt has been made to study the processes involved in the acquisition of verbal repetition habits. Three types of narrative prose material, namely Normal, Ungrammatical, and Inconsistent, were successfully learnt by fifteen subjects. At specified stages during the learning, repetitions were obtained, prompts and corrections being given where necessary, and all errors being recorded and subsequently analysed.

The results appear to indicate that the acquisition of a verbal repetition habit cannot satisfactorily be accounted for on the basis of the older type of trace theory. An alternative interpretation has been put forward based upon Bartlett's conception of 'schema'. On such a view it is suggested that the first stage of learning is the formation of a very general scheme which incorporates a certain amount of dominant detail. Further learning consists in the progressive differentiation and organization of this scheme until finally it becomes capable of mediating literal recall of the material learnt.

The writers would like to express their sincere appreciation to Prof. Bartlett for his many suggestions, and for a great deal of invaluable criticism.

REFERENCES

- (1) BARTLETT, F. C. (1932). *Remembering*. Cambridge University Press.
- (2) BERGSON, H. (1911). *Matter and Memory*. London: George Allen and Unwin, Ltd.
- (3) ENGLISH, H. B., WELBORN, E. L. & KILLIAN, C. D. (1934). "Studies in substance memorization." *J. Gen. Psychol.* xi, 233-60.
- (4) HEYMANS, G. (1913). "Les 'deux mémoires' de M. Bergson." *Ann. Psychol.* xix, 66-74.
- (5) JONES, M. G. & ENGLISH, H. B. (1926). "Notional *v.* rote memory." *Amer. J. Psychol.* xxxvii, 602-3.
- (6) McDUGALL, WM. (1911). *Body and Mind*. London: Methuen.
- (7) SMITH, M. & McDUGALL, WM. (1919). "Some experiments in learning and retention." *This Journal*, x, 199-209.

(*Manuscript received 4 March 1937*)

THE MEASUREMENT OF THE THRESHOLD FOR FLICKER AND ITS VALUE AS A PERSEVERATION TEST

BY S. BIESHEUVEL

*From the Psychological Laboratory of the University of
the Witwatersrand, Johannesburg*

- I. *Theoretical analysis of the flicker test as a measure of perseveration* (pp. 27-29).
- II. *Review of previous work on the threshold for flicker* (pp. 29-32).
- III. *The technique of the flicker test* (pp. 32-34).
 - (1) *The colour of the disk* (pp. 32-33).
 - (2) *Rotation and its measurement* (pp. 33-34).
 - (3) *Position of subject and manner of observing* (p. 34).
- IV. *Results obtained by the application of a new technique* (pp. 34-37).
- V. *Summary* (p. 37).
- References* (p. 38).

I. THEORETICAL ANALYSIS OF THE FLICKER TEST AS A MEASURE OF PERSEVERATION

THE threshold for flicker, produced by the rotation on a colour wheel of two sectors of different colour or of different brightness, has been used in a number of researches as a measure of perseveration.

The assumption is usually made that through the rapid alternation of the two stimuli the actual impulse corresponding to one coincides with an after-effect of the other, producing fusion at a given rotation speed of the disk and flicker below that speed. Gross, Heymans and Wiersma were the first to suggest the existence of such an after-effect, which they called cerebral secondary function. The exact nature of the physiological processes determining flicker or fusion are, however, obscure; though possibly the central nervous function of after-discharge may be responsible for the phenomena. Whatever its final neural explanation, brightness fusion is a particularly good instance of Spearman's law that "Cognitive processes always both begin and cease more gradually than their apparent causes." As such, the threshold for flicker should provide a better measure

of perseveration than the usual ideo-motor tests, which allow of a variety of interpretations by no means all in accordance with the law of inertia.

After a careful analysis of these tests Stephenson⁽¹⁾ concluded that p scores might be the result of four influences:

(1) Continuance and hindrance effects occurring when rapidly switching from one activity to another, or when breaking away from a well-habituated activity or from a newly established mental set. This would be the true psycho-physical perseveration factor in what are usually assumed to be p scores.

(2) Extraneous influence of w characteristics. Unwillingness, laziness, shyness or diffidence can be responsible for the particular efficiency ratio between the two alternating forms of the test obtained. When two children, one keen and interested, the other indifferent, but both fundamentally identical in respect of perseveration, are given a test involving switching from one new activity X to another one Y , both will experience hindrance to the same extent; but the former may have achieved a relatively much higher score in the X activity than the latter, through the energetic way in which he attempted it. Consequently the two p scores would differ considerably, and even a deliberate attempt on the part of the energetic child to overcome the effect of the X activity when doing the Y activity—supposing that to be possible—might not mend matters.

(3) Intrinsic will function. p score may be a measure of the extent to which deliberate volition has overcome continuance effect.

(4) Mental set. If a subject, owing to a large amount of mental inertia, is capable of instituting and retaining a mental set, his performance of a Y activity, normally hindered by the after-effect of a previous X activity, would be facilitated and he would obtain a low p score, when actually he should have had a high one. What are usually taken to be low perseverators should on this interpretation of the tests really be high ones.

Ideo-motor p tests involve furthermore considerable difficulties of technique and scoring, which make interpretation of p scores even more difficult. Such intercorrelations as have been obtained may have been as much due to the constant operation of factors 2 and 3 as to real psycho-physical perseveration.

The flicker test, although it presents technical difficulties of its own, is almost entirely free from the operation of intrinsic and extrinsic will factors. Much less effort is required in observing than in actually performing exacting and fatiguing tasks. The introverted individual will be

much more at ease observing and reporting a subjective experience than performing a task, in which the efficiency of his attempt can be objectively checked.

The extent to which one can, by volitional effort, curtail the after-effect of the type of sensory experience involved in colour or brightness fusion must also be negligible.

Nor could there be any confusion as to whether a low flicker threshold is indicative of high or of low perseveration. A low threshold means a long after-effect and therefore high perseveration. According to Stephenson the facilitating effect of mental set on the *Y* activity only comes into operation when the new task has a "*Gestalt Qualität*" of its own, and is not the summation of units; when the individual can get into the swing of it, and is carried on by virtue of his own inertia. So one can get into the swing of writing S2 S2 S2 S2 S2, the efficiency reached being as high as when writing S or 2 alone. The flicker test is so different in character that it is difficult to imagine how a similar process could take place there, unless there were a tendency for the high perseverator to keep on seeing a flickering surface even at high rotation speed of the motor. As in that case he also would keep on seeing a fused surface once he had developed a mental set for it, this possible flaw could be eliminated by determining the flicker threshold both by means of ascending and of descending series.

The flicker test is therefore superior from a theoretical point of view to the usual ideo-motor *p* tests, and *p* measures obtained from it are far less ambiguous.

Nevertheless, investigators disagree as a result of actual experiment about its value and on the whole are inclined to reject it.

II. REVIEW OF PREVIOUS WORK ON THE THRESHOLD FOR FLICKER

Wiersma(2) found that the threshold for flicker progressively and conclusively differed in groups consisting of manic, normal and melancholic cases. The average speed of the colour disk in rev. per sec. was 27.2 for eleven manic cases; 15.7 for nine normal cases and 12.2 for eighteen melancholic cases. This is consistent with the degree of ideational perseveration characteristic of each group. Other sensory perseveration tests (duration of after effect of a strong electrical stimulus and of a bright light) gave similar results. The smallness of this group was offset by a most careful method of experimentation. The average of 100 threshold determinations, spread over 5 days, was taken. His colour disk with two sectors (red-green) was driven by an electric motor with variable

speed and fitted with a tachometer. In a few manic-depressive cases he found that the threshold for flicker was high in the manic state, low in the depressed state, and intermediate when they were discharged as cured.

The investigation carried out by Lankes (3) differs in respect of method and technique to such a marked degree from the foregoing work that the results are hardly comparable. He used a hand-rotated colour wheel with graded pulleys, such that by turning the handle of a pulley throughout at a constant rate of one rotation per second, and by changing the belt from pulley to pulley, a range of rotation speeds could be obtained varying by steps of 2 rev. per sec. from 25 to 3 rev. per sec. He used three disks with coloured and one with black and white sectors, each of which was observed through a slit illuminated by a 32 candle-power electric lamp, the rest of the room being in darkness. There were four sittings for each sub-group of subjects of from ten to twelve cases, who sat in two arched rows just behind one another in front of the apparatus. At each sitting all four disks were rotated through the complete range of speeds. This was done in a descending series at the first two sittings and in an ascending series at the last two. The subjects were not instructed to watch for fusion or to report flicker; but to identify in the revolving disk one of a range of eighteen colours placed beneath the slit. Marks were assigned according to the correctness of the report, the threshold being calculated from these marks. A different attitude seemed to be involved in the threshold determination by identification of colour than by mere observation of presence or absence of flicker. The black-white disk consistently gave other results than the coloured disks, the correlations obtained between threshold measurements arrived at by the use of this disk and other *p* tests being lowest. Lankes eventually rejected the data from this disk altogether in the final calculation of correlations. Correlating the two trials in either the ascending or descending series with one another showed a high reliability of these measures; but the reliability correlation between the average threshold, obtained from the four disks, both trials ascending, and the same average descending, was only 0.29. In the ascending series subjects started from the stationary position which meant that they knew what colours to look for. In the descending series they started with uniform grey, so that the chances of suggestion influencing the threshold determination were much smaller. Conditions of observation were therefore vastly different in the two series. Lankes did not combine the measurements obtained in the two series, treating those of the descending one as more accurate. Actually, therefore, to Wiersma's 100 threshold

determinations he obtained only two from each subject, which, as they were derived from measurement in only one direction, cannot be taken as accurate. The fairly high reliability of the test merely means a high reliability in measuring whatever the test measures, which may not be the flicker threshold.

Correlations with a number of motor and ideational perseveration tests were for the descending and ascending series respectively (excluding the black and white disk): 0.39, 0.27, 0.40, 0.50, 0.11, -0.05, -0.02 and 0.31, -0.02, -0.02, 0.34, 0.12, 0.27, 0.14, 0.02. With a perseveration questionnaire they were 0.29 and 0.02. The pool of all the tests correlated +0.41 with the questionnaire.

Lankes's procedure has been described at length because it was also adopted by Wynn Jones⁽⁴⁾, when measuring differences in perseveration by means of the flicker test in manic and melancholic cases. He failed to corroborate Wiersma's findings. The following table summarizes the results:

No. of revolutions per minute

	Manic (8)	Normal (25)	Melancholic (5)
Red-green	443	682	541
Blue-yellow	447	674	483

The flicker test failed in this case to distinguish between melancholics and manics in respect of perseveration, although Wynn Jones himself had found that on the basis of motor tests there was a slight difference between them. Using similar tests, Stephenson⁽⁵⁾ had found melancholics to be high perseverators and manics non-perseverators, provided there was no supervening dementia. Clarke⁽⁶⁾ included a flicker test in a battery of *p* tests applied to delinquent and normal children. He used a black and white disk on a colour comparator fitted with a tachometer measuring up to 4000 rev. per min. The disk, which was illuminated by an electric light of high candle-power, was viewed through a stereoscopic eyepiece without lenses, in order to lessen the effect of head and eye movements. Subjects were instructed to report appearance and disappearance of flicker in descending and ascending series. Each series was repeated three times and the average of the six readings obtained was taken as the threshold. Subjects also reported two dark revolving sectors on the disk, which at certain speeds appeared to be stationary. The average of six speeds at which this occurred was also noted for each of the fifty subjects. Although these two measures obtained from the flicker test intercorrelated 0.72 ± 0.0455 , their correlation coefficients with a number

of motor and ideo-motor tests of perseveration were, all but one, insignificant. The exception was a correlation between the real flicker measure and a switching (alphabet-number) test (0.349 ± 0.09), which itself showed a very poor average intercorrelation with the others (0.196 ± 0.09). Clarke concluded that the flicker test appeared of doubtful value as a measure of perseveration, "unless there is a large chance factor entering into the measurement of the limen for flicker—and the high correlation between 6*a* (true flicker) and 6*b* (stationary sectors) seems to preclude this".

III. THE TECHNIQUE OF THE FLICKER TEST

Where there is so much divergence, not only in what is being measured but also in the way of measuring it, it is hardly surprising that results have been in most cases unconvincing and that on the whole ideo-motor tests with all their failings have been preferred as measures of *p*.

The determination of the threshold for flicker presents a number of problems of measurement. In the first place, the observations required from the subject are by no means easily made. Secondly, small threshold differences are of the utmost importance, and their accurate recording is essential. Apart from that, numerous extraneous factors can vitiate the threshold measurement. Unless a careful and standard technique is used, it is idle to discuss the practical value of the test. I shall deal in turn with the various factors determining the measures obtained.

(1) *The colour of the disk*

A disk with black and white sectors is preferable to one with coloured ones. Colour adds an unnecessary and complicating factor. Flicker produced by the rotation of the black-white disk consists only of the element of coming and going, apparent movement observed against a grey background. The subject's attention can be directed towards that. Flicker produced by rotation of the colour disks involves a qualitative comparison between the ultimate grey of fusion and the alternating colours prior to it. The subject is not so much observing the presence of flicker, alternation, as trying to find traces of two colours in a grey background. It is easier to imagine one still sees two colours in a grey, which as a sensory datum belongs to the same class as colours, than to imagine one sees in the grey surface a flickering movement, which belongs to a quite different class of visual data. The instructions given by Lankes (3) to his subjects to match the colours seen at any given speed with those provided as standards would of course put the black-white disk at a great disadvantage. Red and green, though very indistinct at medium speeds, remain qualitatively visible up to fusion when rotated; whereas rotation of the black-white disk almost immediately produces an alternation of bright and dark greys. The black and white can hardly be qualitatively recognized at low speeds, and eventually the bright and dark greys can only be separated from each other on the ground of flicker. This accounts for the discrepancy between the results which Lankes obtained from the black-white disk and those obtained from the colour disks.

It is therefore recommended that a disk be used, one-half of which is black and one-half white. This disk should be observed through a circular aperture, about 1 in. in diameter. The disk must be artificially illuminated in order to secure constancy of brightness.

(2) *Rotation and its measurement*

Although normal subjects differ considerably in respect of the presence or absence of perseveration qualities in their everyday behaviour, such behavioural differences may depend on very minute differences in physiological function, and hence also on equally minute differences in sensory function. Apparatus is therefore required which allows of very gradual changes in the speed of the flicker disk and an accurate measurement of these changes. For this purpose, Lankes's apparatus is wholly inadequate, not only in that the steps between the successive speeds are far too big, but also because hand rotation, no matter how carefully performed, is bound to involve irregularities.

After a good deal of experimentation, the following apparatus was eventually designed:

An electric induction motor (A.C.) of 1/40th H.P. was geared at right angles to a light metal disk similar to a gramophone turntable, by means of a worm and pinion on its spindle. On the other end of the spindle an aluminium disk, one-half black and one-half white was mounted. The current was supplied to the motor by means of a double transformer. By using half or the whole of this transformer and a large number of tappings operated by a switch, the voltage could be controlled so as to give a range from very low to very high speeds. A small dial rheostat covered the difference in supply between two tappings. A magnet fixed around the metal flicker disk in front put additional load on the motor, whereby very low speeds could be obtained and irregularities in the running could be smoothed out. Hence a continuous unbroken increase in speed was possible, which could be very accurately regulated by means of minute alterations in voltage.

Good tachometers are expensive and put additional load on the motor. Instead, a device similar in principle to the familiar stroboscope used to set a gramophone to the required speed was used. A Neon lamp was fixed above the turntable of the motor, which carried a cardboard disk on which a number of concentric rings were drawn. Each of the resulting concentric rings was divided by radial lines into a large number of equal sections. The angle between these radial lines was increased with each successive circle, its size being determined by the following facts: For every ring constructed in this manner there is a particular speed such that, when it is viewed under the light of a Neon lamp, the radial lines dividing the ring in question appear to be standing still. With fifty-cycle alternating current, a Neon lamp has fifty periods of maximum and fifty periods of minimum brightness. A period of minimum brightness therefore lasts $\frac{1}{100}$ sec.

Consequently if one wants a ring to be stationary at a speed of 1 rev. per sec. of the turntable (i.e. 20 rev. per sec. of the flicker disk) the radial lines must be spaced in such a way that each moves through as many degrees as separate it from its neighbour in the period of minimum illumination ($\frac{1}{100}$ sec.). In this case the angle between the radial lines would therefore be 3.6 degrees. The stroboscope actually used was constructed to measure a wide range of speeds, well beyond the usual distribution of

thresholds. In the range of the most commonly occurring threshold values, successive rings measured speeds differing from each other only at the rate of 1 rev. per sec. When two adjacent rings move in opposite directions at approximately equal speeds, the speed of the flicker disk will be midway between the speeds represented by the two rings when apparently stationary. By assessing the relative rates at which two rings move in opposite directions, quarters can easily be interpolated, the real speed always being nearer to that indicated by the one going slower. By means of such interpolations between rings, it was possible to get very fine readings, correct to a fraction of a revolution per second. Smooth working of the apparatus was essential to ensure such accuracy, and in this respect our method of transmission could still be considerably improved.

The set-up as described enabled the experimenter to read off, simultaneously with the subject's report, the speed at which it occurred.

(3) *Position of subject and manner of observing*

The distance of the subject from the flicker disk and the angle from which he views it both affect the threshold. When no flicker is observable when the disk is looked at from a position at right angles to it, it frequently returns when one looks at it from the side. Flicker also persists much longer at the periphery of the eye than at the fovea. Sudden head or eye movements, especially winking, will also bring the flicker back. The threshold is furthermore affected by the dark adaptation of the eye. The room must also be reasonably quiet, for it has been shown that harmonious and rhythmical sound favours the disappearance of flicker, whereas noise and dissonance retards fusion. Finally fatigue due to prolonged fixation tends to make observations unreliable. Lankes's subjects were seated at different distances from the disk and observed it at different angles; the session lasted 40 min., most of which time the subjects were observing. Such conditions precluded the obtaining of any valid data. The position of the subject should therefore always be immediately in front of the disk and at a constant distance from it. No observations should be started before the subject is dark adapted. The light illuminating the disk may only shine while the subject is observing; each observation must start on a signal from the experimenter and last only for a few seconds. Head movements should be guarded against. Clearly most previous investigators did not adequately provide for these sources of error in their technique and the failure of the flicker test as a measure of p may be due to this, rather than to any lack of connexion between the flicker threshold and perseveration.

IV. RESULTS OBTAINED BY THE APPLICATION OF A NEW TECHNIQUE

As criteria for perseveration or non-perseveration a number of items in a questionnaire, used by the writer⁽⁷⁾ in an investigation into the nature of temperament, were selected. A group of 645 children were rated on these questionnaire items, receiving a + mark or a - mark when they possessed what was considered to be a perseveration or a non-perseveration trait, respectively. According as their final score was + or - they were classified as perseverators or non-perseverators. An exami-

nation of the percentage incidence of qualities in the two groups so obtained revealed that the perseveration and non-perseveration criteria respectively cohered and were the expression of the opposite aspects of one underlying factor. These findings are summarized in Table I, which also gives the weights assigned to each criterion on the basis of some tentative preliminary trials and evidence of their diagnostic value obtained from the work of others in this field. The reliability of the difference between incidence of each trait in one group and in the other group is also indicated.

Table I. *Selective value and cohesion of perseveration and non-perseveration criteria. Percentage incidence of each criterion quality in a group of 310 cases selected for perseveration and in a group of 335 cases selected for non-perseveration*

Criteria	Weight	% incidence in perseverating group (310 cases)	% incidence in non- perseverating group (335 cases)	% incidence difference between perseverating and non- perseverating groups	Standard deviation of difference
Works regularly at constant rate	P 1	79.0	16.4	62.6	3.1
Works irregularly	N -3	14.2	70.8	-56.6	3.2
Usually attentive	P 2	94.0	22.4	71.6	2.6
Easily distracted	N -3	3.5	48.4	-44.9	2.9
Playful	N -3	1.2	23.0	-21.8	2.4
Busy with other things than lesson	N -3	3.1	25.0	-21.9	2.5
Stops attending at bellring	N -2	6.2	68.0	-61.8	2.9
Retains previously learned task	P 1	61.1	34.7	26.4	3.7
"Scatterbrained"	N -1	23.8	51.8	-28.0	3.7
Retains when interested	N -1	10.4	12.5	- 2.1	2.4
Impulsive	N -2	11.0	60.6	-49.6	3.2
Cautious	P 2	89.0	38.1	50.9	3.2
Cheerful	N -1	33.6	49.1	-15.5	3.8
Depressed	P 1½	11.5	6.2	5.3	2.2
Alternates from one mood to another	N -3	11.7	39.0	-27.3	3.2
Always even-tempered	P 3	46.6	10.8	35.8	3.3
Constant in friendship	P 1	78.3	51.2	27.1	3.6
Inclined to be fickle	N -1	5.8	30.6	-24.8	2.8
Restless, inclined to fidget	N -2	6.4	68.0	-61.6	2.9
Calm and quiet	P 1½	72.2	19.1	53.1	3.3
Av. % incidence perseveration criteria		66.5	24.9	41.6	3.6
Av. % incidence non-perseveration criteria		10.9	45.6	-34.7	3.2

Two groups of schoolboys, numbering nineteen and twenty-three cases respectively, were then rated on these criteria by their respective class masters. I was unfortunately unable to get two ratings for each boy, so that all that one can say of the resulting groups is that on the average perseverators predominate in the perseveration group, non-perseverators in the non-perseveration group. Doubtful cases were not rejected but classified one way or another. These boys were then tested with the apparatus and technique described.

The data from the two groups of nineteen and twenty-three cases were not combined, since the flicker apparatus, though identical in both cases, was set up in different localities. Both the subjects' distance from and the illumination of the disk differed on the two occasions.

Each subject was tested individually. First he was shown a very obvious flicker. Thereafter the motor was speeded up, so that there was no possibility of flicker at all. This was done in order to acquaint him with the different situations to which he had to react. The subject was then told: "Whenever I ask you to do so, look straight and without moving at the disk. As soon as you can see the flicker say 'Yes', or if you cannot see it say 'No'. Make quite certain whether you can or cannot see it. Look down in front of you when not reporting. Only look at the disk when asked to do so." As soon as the subject was ready to report, a 4 V. lamp at fixed distance from the disk was switched on. Though not strong enough to cause any dazzle, the lamp threw a sufficiently clear beam on the disk. The subject was shielded from its direct light.

Six threshold measurements were then taken, three increasing from very low speeds, three decreasing from high speeds. Between the rough limits so established, frequently not much more than 1 rev. per sec. apart, about thirty further readings were taken—often more, sometimes less—depending on the steadiness of the subjects. The threshold was taken to lie between the two most frequent limiting positions. This method prevented fatigue due to frequent repetition of long series. Only the preliminary threshold determinations demanded continuous observation. After that the subject merely momentarily fixated the disc and reported on absence or presence of flicker. By this method also there was less likelihood of the subjects' habitually reacting at a certain point in the series. And finally most measures were taken in the immediate neighbourhood of the threshold so that by means of small changes in speed its position could be very accurately determined.

In each group the average flicker threshold for the perseverators and non-perseverators was calculated, and the standard deviation of the difference between them determined (see Table II).

Both these differences, though small, are significant, the one being three times, the other four times, its standard deviation. As both are in

Table II

	Perseverators	Non-perseverators	Difference	Standard deviation of difference
Group A	35.3	37.5	2.2	0.84
Group B	31.6	33.1	1.5	0.386

the same direction, the conclusion is warranted that behavioural perseveration and sensory perseveration, measured in terms of the threshold for flicker, tend to vary together; and hence that the flicker test, when properly standardised, can be a useful measure of perseveration. The implications of Wiersma's results⁽²⁾ obtained from mental patients have therefore been shown to hold for normal subjects. Wynn Jones's contradictory evidence⁽⁴⁾ must be rejected, his apparatus and technique not being suitable for the measurement of the threshold for flicker.

Whether the flicker test will correlate with the ideomotor tests remains to be seen. The technique of the latter still presents considerable difficulty, and such intercorrelations as have been obtained between them may in part be due to the operation of those intrinsic and extrinsic will factors mentioned by Stephenson⁽⁵⁾. Pinard⁽⁸⁾ also found it necessary to mention this possibility when discussing his results.

V. SUMMARY

Though possessing considerable theoretical merits, the threshold for flicker proved in practice to be a poor measure of p .

Previous investigators differed extensively, however, in their manner of measuring this threshold, most methods being subject to serious errors. The measures obtained by them tended to be functions of their own technique, rather than of p .

Apparatus and technique were therefore worked out which took into account the many factors which were found to affect the threshold for flicker, or which were peculiar to its measurement.

When this new method was tried out on two groups of subjects, previously rated on the presence or absence of perseveration, it was found that there was a significant difference in respect of the flicker threshold between the two groups.

Sensory perseveration, found to be connected with general behavioural perseveration qualities, can therefore be measured by means of the flicker test.

REFERENCES

- (1) STEPHENSON, W. (1934). "An introduction to so-called motor perseveration tests." *Brit. J. Educ. Psychol.* iv, 186.
- (2) WIERSMA, E. (1906). "Die Sekundär Funktion bei Psychosen." *J. Psychol. Neurol., Lpz.*, viii.
- (3) LANKES, W. (1915). "Perseveration." *Brit. J. Psychol.* vii, 387.
- (4) WYNN JONES, LL. (1929). "Individual differences in mental inertia." *J. Nat. Inst. Industr. Psychol.* iv, 282.
- (5) STEPHENSON, W. (1932). "Studies in experimental psychiatry, II." *J. Ment. Sci.* lxxviii, 315.
- (6) CLARKE, G. (1934). *Some Character Traits of Delinquent and Normal Children in Terms of p factor*. Melbourne University Press.
- (7) BIESHEUVEL, S. (1936). "Thé nature of temperament." *Trans. Roy. Soc. S. Afr.* xxiii, 311.
- (8) PINARD, J. (1933). "Tests of perseveration." *Brit. J. Psychol.* xxiii, 5, 114.

(*Manuscript received 10 January 1937*)

THE EFFECT OF KNOWLEDGE OF RESULTS ON LEARNING AND PERFORMANCE

I. A CO-ORDINATED MOVEMENT OF THE TWO HANDS

BY J. L. ELWELL AND G. C. GRINDLEY

From the Psychological Laboratory, Cambridge

- I. *Introduction* (pp. 39-40).
- II. *Apparatus* (pp. 40-42).
- III. *Experimental procedure* (pp. 42-43).
- IV. *Results* (pp. 43-50).
 - (1) *Performance without knowledge of results* (pp. 43-44).
 - (2) *Performance with knowledge of results* (pp. 44-49).
 - (3) *Removal of knowledge of results* (pp. 49-50).
- V. *Discussion* (pp. 50-53).
- VI. *Summary* (p. 53).
- References* (p. 54).

I. INTRODUCTION

A NUMBER of experiments have been reported by various writers in which knowledge by the subject of the success or failure of his performance has been used as an incentive. For instance, the experiments of Arps (1917) on ergographic work, of Crawley (1926) upon work done with the arm and leg muscles, and of Johanson (1922) upon reaction times have shown that in all these tasks knowledge of results can produce an improvement in performance. In many of Thorndike's recent experiments on human learning (Thorndike, 1931, 1932, 1935) the subject has been placed in a situation in which he can make any one of a number of discrete responses, and has then been told whether his response is 'right' or 'wrong'; and Thorndike has studied in detail the way in which knowledge that a response is 'right' leads to an increased tendency to perform that response in future. In these experiments the word 'right' seems to act in very much the same way as a 'reward' is supposed to act in experiments on learning in animals. In the other experiments which have been referred to, knowledge of results seems to act in rather a different way—it seems to act as an incentive to greater effort rather than as a selective

40 *Knowledge of Results on Learning and Performance*

influence leading to one specific response in preference to another. The experimental studies so far published are sufficient to show the importance of knowledge of results as a factor in learning and performance, and they suggest that knowledge of results may operate in several different ways; but it is clear that there is need for further investigation.

The experiments described in this paper were designed as a detailed study of the way in which knowledge of results operates in the acquisition and maintenance of a skilled muscular performance, and as an investigation of the effect of removal of knowledge of results after a skilled habit has been acquired. The performance which we studied was a relatively simple one—it involved a co-ordinated movement of the two hands so as to direct a spot of light on to the ‘bull’s-eye’ of a target, the inaccuracy of the extent of movement by either hand being registered by the deviation of the position of the spot of light from the bull’s-eye in a corresponding direction on the target. (Thus, too small a movement by the subject’s left hand caused the spot to be too far ‘up’ on the target, too small a movement by the right hand caused the spot to be too far to the left on the target, and so on.) The position of the spot of light normally gave the subject knowledge of his results (i.e. of his score, and of the extent and direction of any error which he had made); but it was also possible to arrange the apparatus so that the subject did not see what position on the target he had reached, and so had no knowledge of his results.

II. APPARATUS

The apparatus, which is shown in Fig. 1, consisted of a wooden frame *AAAA* measuring $2\frac{1}{2} \times 2\frac{1}{2} \times 1$ ft., open at the sides, and the subject stood facing this frame at the point indicated by the arrow *B*. Inside the frame was a system of steel arms *CDEFGHI*, hinged together at all these points, the whole system being supported by a strong bearing at *E*. A spring-loaded pencil *P* made a record on a sheet of graph paper below it of the form of any movement of the system of arms, and a torch *T* was arranged so that, when switched on, it threw a small spot of light on to a ground-glass target *JJ*. Thus, if the handle *K* was held still by the subject, movement of the handle *L* caused the spot of light to move (in an arc of a circle) to the subject’s right or left; and, if *L* was held, movement of *K* caused the spot to move, in an arc of a circle, towards or away from the subject. The dimensions were arranged, for convenience, so that the movements of the pencil *P* were geometrically similar to, and twice as large as, the corresponding movements of the torch *T*. The target *JJ* was 2 in. in diameter and consisted of ten concentric rings

increasing in diameter by equal steps of 0.2 in. (In the actual experiments a score of 10 was given if the handles were moved so that the spot appeared in the bull's-eye, 9 for the next ring, and so on, so that the outer ring scored 1, and if the target was missed altogether the score was zero.) Rubber stops, not shown in diagram, limited the movement of the arms so that the spot could not be moved more than a few inches from the target in any direction. The whole apparatus moved silently and smoothly, so that the subject could not infer the position of the arms from any noise or frictional resistance to the movement.

Two tapping keys (not shown in Fig. 1) were placed below the handles *K* and *L*, and in all the experiments the subject was instructed to press these keys before and after making the required movement with the

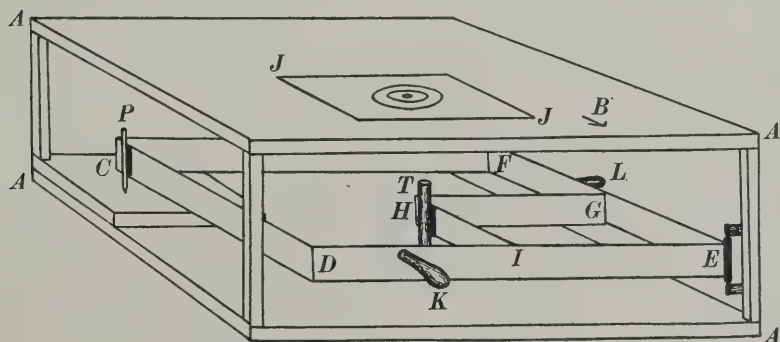


Fig. 1.

handles. An electrical circuit, with a switchboard controlled by the experimenter, made it possible to produce any of the following arrangements of the apparatus:

(a) *For practice trials, to demonstrate the principles of the apparatus*, the torch remained switched on all the time, and the subject was thus able to study the effect of moving each arm of the apparatus.

(b) *For experiments with knowledge of results*, the torch was on when the subject pressed both keys, and off when he released them. Thus the subject saw the spot of light at its starting position, could not see it while he was making the movement, but saw it again immediately he returned his hands to the keys after making the movement. A chronometer in the circuit registered the time interval between the moment at which the subject released the keys and the moment at which he depressed them again after making the movement. (Since the time taken

42 *Knowledge of Results on Learning and Performance*

to move from the keys to the handles *KL* and back again after the movement of the arms had been made became more or less standardized during the preliminary training, the chronometer time gave an indication of the time taken by the subject to make the required movement of the handles.)

(c) *For experiments without knowledge of results*, the torch remained off all the time, and the subject was therefore unable to see what part of the target he had reached. The chronometer recorded the time of the movement as in arrangement (b).

III. EXPERIMENTAL PROCEDURE

Preliminary training. Before experimenting for the first time the principle of the apparatus was explained to each subject, and, with arrangement (a) of the circuit, the subject was allowed to move the handles and to observe the resulting movement of the spot. The subject was *not* allowed to move the spot of light from the starting point to the bull's-eye during this training, i.e. he was not allowed to carry out the actual movement which he would be required to perform in the subsequent experiments.

Performance with knowledge of results. Before each trial the experimenter placed the two arms of the apparatus in a fixed position which was such that the spot of light was about 2 in. from the target, in such a direction that the subject had to pull both handles about the same amount in order to reach the bull's-eye. The subject was given the following written instructions, which could be referred to by him during the experiment:

When the experimenter says 'Ready' place your hands on the keys and depress both keys.

When the experimenter says 'Go' take hold of the handles and move the arms to the position which you judge to correspond to the 'bull's-eye' of the target, and then immediately return your hands to the keys and depress both. In making the movement aim primarily at accuracy. The speed of the movement will also be taken into account in marking the results; but do not sacrifice accuracy to speed.

Keep the keys depressed until the experimenter says 'Right'. Then read out your score.

The number and spacing of the trials carried out at each sitting varied in different experiments, and will be described with the results.

Performance without knowledge of results. The procedure was exactly the same as for experiments with knowledge of results; excepting that

the subject was told that he would not see the light, and that arrangement (c) of the circuit was used instead of arrangement (b).

In all the experiments, both with and without knowledge of results, the pencil record on the graph paper gave the experimenter a record of the form of the movement made by the subject, which showed whether he had pulled both handles together, or one before the other, or had executed some more complicated movement. The chronometer time gave an approximate record of the speed at which the movement had been made.

The subjects were also asked, at intervals, for introspective reports; since it was hoped to obtain from these further light upon the process of learning a task of this kind and upon the effects of knowledge of results. Introspective reports were only asked for after every ten or twenty trials (as is described in the next section), since it was feared that the attempt to introspect after every single trial would interrupt the continuity of the learning process. In order to avoid any danger of suggestion by the experimenter of what the subject was expected to say, the request for an introspective report was always put in the form "Have you anything to say about that?" or "Can you give any introspections?"; and no further questions were asked except in neutral forms such as "Is there anything more you can say about it?" The reports of the subjects were taken down by the experimenter. (The disadvantages of this 'natural history' method of taking introspections are obvious, the main ones being that it tends to produce reports which are very brief and unsystematic, and which often contain no reference to the aspects of the process in which the experimenter is most interested. Nevertheless, in an investigation of the kind described here, the great advantage of relative freedom from the danger of suggestion by the experimenter seems to outweigh the disadvantages.)

IV. RESULTS

(1) *Performance without knowledge of results*

It is clearly not to be expected that in an experiment of this kind performance without knowledge of results would show any consistent tendency to improve in successive trials (cf. Thorndike, 1932, pp. 8-15). Nevertheless we made a few experiments to test this. After preliminary training five subjects were each given a series of fifty trials without knowledge of results at one sitting. The scores of each subject showed a wide scatter, the target being missed altogether in many of the trials, and the subject's accuracy showed no consistent tendency either to

44 *Knowledge of Results on Learning and Performance*

increase or to decrease. The times occupied by the movement showed a tendency to decrease, the subjects tending to make simpler and quicker movements as they became more 'bored' with the experiment. They were asked afterwards if they had any introspections to offer; and they all said that they felt no confidence in their performance, and expressed, in various ways, their distaste for the experiment.

These experiments are too slight to be worth reporting in detail; but they serve as a check on the fact that the improvement of performance with knowledge of results (described in the next paragraph) was really due to the visual clue given by the spot of light on the target, and not to any other clue given by the apparatus or by the experimenter. Further confirmation of this is given by the fact that when four of these five subjects were afterwards given trials with knowledge of results their performance showed a marked improvement within the first ten trials. They reported that the experiment had now become 'interesting' or 'more enjoyable'; and they now took more care over making the movement, so their time of response became longer.

(2) *Performance with knowledge of results*

Quantitative results. A number of experiments were made to study the way in which performance improves with practice when knowledge of results is given. Only two of these need be reported in detail.

In Exp. I ten subjects (all of them women students, between the ages of 18 and 22, and not students of psychology) were each given twenty trials in immediate succession on each of ten successive evenings. A series of twenty trials occupied on an average about 8 min., and at the end of the series the subject was asked if she had anything to say. In Fig. 2, *AB* is the average curve of learning of the ten subjects, each point representing the mean of ten trials for all the subjects (i.e. each point representing the mean of 100 readings). Nine of the ten subjects began with low average scores and improved progressively; the tenth subject, who began with an abnormally high score (an average of 7.6 in the first ten trials), showed little or no further improvement.

In Exp. II ten subjects (all students or research workers in experimental psychology) were given ninety trials at one sitting. Rest pauses of a few minutes were allowed after the twentieth, fortieth, sixtieth and eightieth trials, and in these pauses the subjects were asked for introspections. *AB* in Fig. 3 is an average curve plotted in the same way as *AB* in Fig. 2. All the subjects showed a progressive improvement in score.

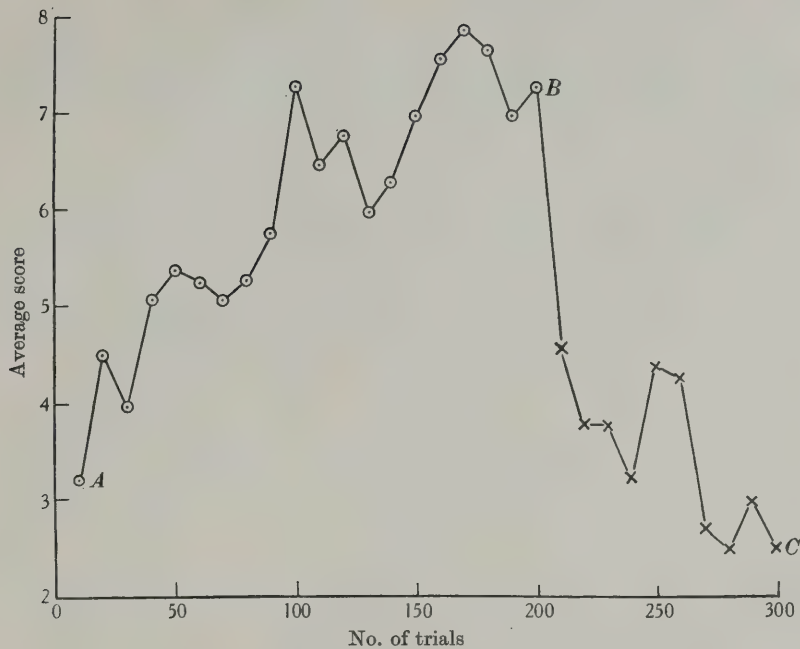
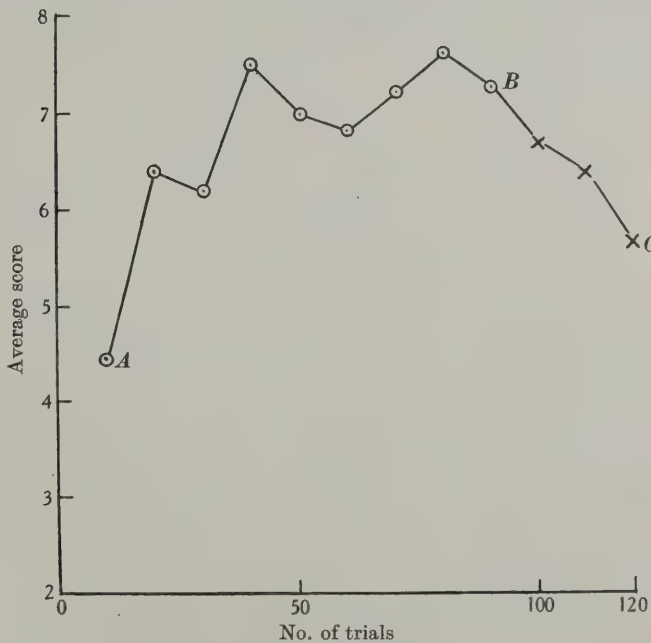


Fig. 2. Average of performance of 10 subjects in Exp. I. (Knowledge of results removed at B.)



46 *Knowledge of Results on Learning and Performance*

In both experiments the actual scores in individual trials were very irregular. Fig. 4, which is a typical part of a detailed learning curve, shows the scores of one subject (in Exp. I) on the three first days of learning.

Objective characteristics of performance. Most of the subjects in both the experiments described, while showing an improvement in the accuracy of their performance, showed a slight decrease in the time taken to perform the movement. This decrease in time (about 25% on the average between the first ten trials and the last ten trials in Exp. I)

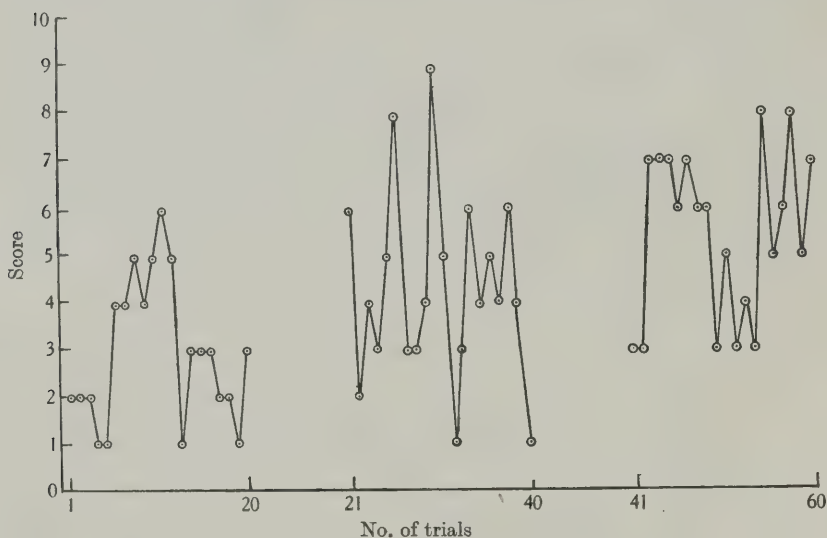


Fig. 4. Individual scores of one subject during the first three days of learning in Exp. I.

seemed to be attributable to the fact that they hesitated less as the training proceeded. Much more striking than this, however, were the persistent differences in time between one subject and another. Table I, which is typical, shows the time taken by four subjects in ten successive trials (trials 7-16) near the beginning of the learning process. One subject, O, always seized both handles as soon as the experimenter said 'Go', and made a rapid pulling movement, so that his time of response was never more than 2.2 sec.; while another subject, H, always pulled the handles very gently, so that her time of response was never less than 12.7 sec. These differences nearly always persisted throughout the learning process, apart from some aberrations during the first few trials.

Table I. *Times of movement (in seconds) of four subjects near the beginning of the learning period (Trials 7-16)*

Trial	Subject			
	O	Z	V	H
7	1.4	4.8	7.3	16.4
8	1.5	3.5	6.7	20.3
9	1.9	3.5	6.2	12.7
10	2.2	2.9	7.0	13.1
11	1.4	3.5	7.5	12.9
12	1.4	3.1	7.3	16.2
13	1.3	3.4	8.2	14.5
14	1.3	2.5	7.6	16.5
15	1.6	2.8	6.6	13.4
16	1.1	2.4	8.3	12.8

Not only the time taken by the movement, but also the form of the movement (as recorded by the pencil) showed characteristic differences between the subjects which usually persisted throughout the learning, or for a large part of the learning. Fig. 5 shows tracings of the forms of four successive movements by each of four subjects in the middle of the learning period. Subject D, it will be seen, moved one arm of the

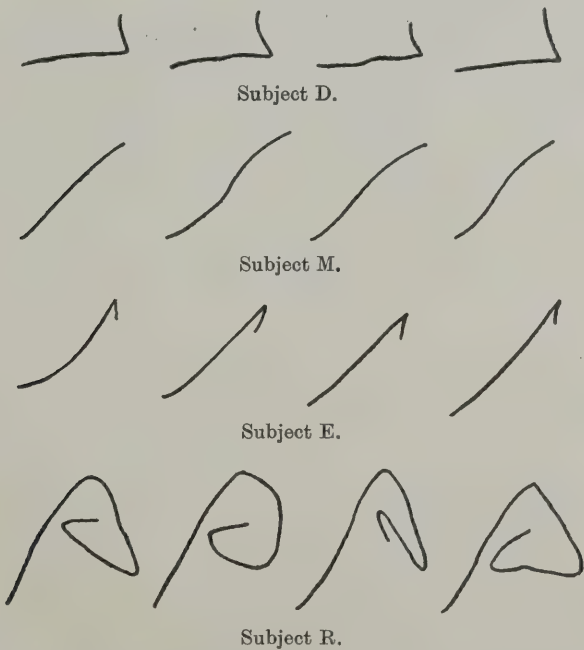


Fig. 5. Tracings of pencil records, showing the form of successive movements by four subjects.

48 *Knowledge of Results on Learning and Performance*

apparatus before the other, while Subject M moved both arms simultaneously, and the other two subjects made somewhat more complicated movements. These stereotyped patterns of response usually developed early in the learning process and remained constant; but there were a few subjects who made abrupt changes in their pattern of response during the experiment. When these changes occurred they were nearly always described (with reasons) in the introspections of the subject.

Introspective reports, and general attitudes of the subjects. In Exp. I the subjects all appeared keenly interested in the task, and they nearly all frequently expressed satisfaction at a high score and disappointment at a low score. In Exp. II the subjects, who were all students of experimental psychology and had therefore done experiments of this general kind before, showed rather less enthusiasm; but even with these subjects expressions of satisfaction or dissatisfaction were frequent. Expressions of dissatisfaction at a low score were most frequent when this followed a series of high scores. When a bull's-eye was obtained the subjects often made remarks such as, "I must try to do exactly that movement again". When the bull's-eye was not hit several subjects reported that, early in the experiment, they had to formulate to themselves in words such ideas as, "That means that I must pull farther with the left hand", while other subjects reported that the tendency to correct for an error made in the previous trial was "conscious, but more or less automatic". Although the reports about this tendency to correct for errors are relatively few it seems probable from them that at first, when the apparatus was unfamiliar, many of the subjects found it necessary to think out explicitly the type of correction called for by any direction of error on the target, but that as the apparatus became more familiar this correction became more automatic, and explicit formulation no longer occurred.

The persistent differences between different subjects in chronometer time and in the form of the movement have already been mentioned. It appeared from the introspective reports that these differences usually corresponded to differences in the methods which the subjects consciously adopted in order to carry out the task. Thus some subjects reported that they always tried to pull both handles equally hard for a given length of time at a speed which they tried to keep constant, the correct time being judged by some device such as counting five (subvocally); other subjects carried out similar calculations but always moved one arm before the other; others always attempted a rapid movement and tried to regulate it by the 'amount of force' applied to

each handle; and still others decided that they could be most accurate by carrying out one or other of these schemes and then applying a subsequent 'correction' movement to compensate for a probable error in the original movement. In a few cases it was possible to explain why a subject adopted a particular type of movement. For instance, one subject reported that during one of the early trials, having moved both arms simultaneously, he felt that he had gone too far to one side, and to correct for this he made a small movement of one arm only. The result was a bull's-eye, and the subject decided always to move the arms in this way, and so developed a pattern of movement similar to that shown for Subject E in Fig. 5. But the reasons why one subject chose one method and another chose another method did not usually appear from the introspective reports. Other differences of conscious attitude which the subjects reported were not reflected in any clear differences in objective performance. For instance, some subjects reported that when making the movement they visualized the spot of light crossing the target, and they continued to move until this imaged spot reached the bull's-eye; while others reported that they made their judgements entirely on the basis of kinaesthetic sensations (e.g. that they moved the arms until they "got the right feel from the wrists").

(3) *Removal of knowledge of results*

In Exp. I, which has been described in the previous section, the conditions were changed so as to give no knowledge of results after the 200th trial (i.e. at the beginning of the eleventh day), and the experiment was continued for 100 more trials (i.e. for five more days). The average performance of the subjects (which is represented by *BC* in Fig. 2) showed an immediate and very marked deterioration. Only two of the subjects maintained a level of performance near to their previous level for as many as forty trials after the change of conditions. The earlier pattern of movement was usually retained; and the scores of the subjects in individual trials continued to show a wide scatter. All subjects expressed displeasure and annoyance at the change of conditions, and most of them said that they now felt no confidence in their results. Their dislike of the new conditions was shown also by a general attitude of boredom in place of their previous attitude of keenness. (The subjects also began to arrive late for the experiment, and often asked how many more days it would last.) As the subjects became more bored with the experiment they appeared to become careless in making the movement, and with

50 *Knowledge of Results on Learning and Performance*

most of them the chronometer times showed a corresponding tendency to decrease.

In Exp. II a similar change in conditions was made after the ninetyeth trial, without any pause longer than that required to tell the subject that he would no longer see the light, but must continue to aim at the bull's-eye. Thirty trials were given without knowledge of results, with the usual rest pause (during which introspections were taken), after the 100th trial. The average performance of the subjects is represented by *BC* in Fig. 3. Nine out of the ten subjects had a lower average level of performance in the thirty trials without knowledge of results than in the previous thirty trials; the tenth subject had a very slightly higher average level for the trials without knowledge of results. The subjects, who were all more or less accustomed to monotonous experiments of this kind, did not show such a marked emotional reaction to the change of conditions as did the subjects in Exp. I. But most of them said that they now found it 'boring', or that it was now difficult to prevent their thoughts from wandering. Several subjects explained that they were attempting to do exactly as before, but that it was difficult to remember what a correct movement felt like now that there was no 'check' on their performance.

It will be seen that the average deterioration of performance after removal of knowledge of results is much more rapid in Exp. I than in Exp. II. Comparisons between the records of the subjects in these experiments, and comparison of these results with the results of some similar experiment, which need not be reported, suggest that this is probably due to two factors, (*a*) the greater boredom and annoyance of the subject in Exp. I, and (*b*) the fact that in Exp. I there was an interval of 24 hr. before the first trials without knowledge of results, while in Exp. II these trials followed almost immediately the last trials with knowledge of results. We did not attempt any quantitative estimate of the relative importance of these two factors in determining the rate of deterioration of performance after removal of knowledge of results.

V. DISCUSSION

The quantitative results described in this paper suggest, at first sight, that 'knowledge of results' in the acquisition of a human skill may be regarded as acting similarly to a 'reward' in animal learning. If, for instance, we compare the results described here with the results of experiments on the learning of mazes and other tasks by animals, we

notice at once that there is no appreciable learning when no knowledge of results is given or when no reward is given (cf. Blodgett, 1929); that learning occurs when knowledge of results is given or when a reward is given (cf. Tolman, 1932; Grindley, 1932); and that the acquired habit breaks up when knowledge of results is removed or when the reward is removed (cf. Bruce, 1930; Grindley, 1932). There is also an obvious parallel between the subject's attitude (e.g. of 'keenness' or 'desire to do well') in this experiment and the 'drive' in the animal studies (cf. Tolman, 1932).

It is therefore tempting to try to explain the results of these experiments in terms of principles, such as Thorndike's 'Law of Effect', which have been used to explain animal learning. But in the acquisition of a muscular skill, such as that described in this paper, the learning cannot be regarded merely as the strengthening of the tendency to repeat movements which have been 'rewarded' (by a high score). If a subject missed the bull's-eye he tried, next time, to correct for his error by altering his response in the appropriate direction.

A similar attempt to correct for errors obviously occurs in many other skilled performances, e.g. in ball games. In many of Thorndike's experiments (1931, 1932, 1935), where the subject is allowed to vary his behaviour only between a limited number of discrete responses, and is told whether he has made the 'right' response or a 'wrong' response, it may be legitimate to consider learning simply as the strengthening or weakening of the tendencies to make each of these responses. But in experiments such as the present ones we have to consider that knowledge of results, when the movement was not completely successful (i.e. when it did not result in a bull's-eye) introduces also *a tendency towards a specific kind of variation of the response which has just been made*. We may call this the 'directive effect' of knowledge of results. It seems clear that principles such as the Law of Effect would require considerable modification in order to be able to explain satisfactorily any form of learning in which the 'directive effect' is important. We shall not attempt here any comprehensive statement of a satisfactory principle of learning applicable to muscular skill. Further analysis of the 'directive effect' is required before this can usefully be done.

There appear to be three main ways in which knowledge of results leads to improvement of performance in the experiments described in this paper:

(a) By causing a tendency to repeat movements which lead to a satisfactory result. This can be regarded as an example of the operation

52 *Knowledge of Results on Learning and Performance*

of the Law of Effect. The tendency, it should be noticed, is sometimes fully conscious under the conditions of these experiments; and the subject may formulate to himself exactly what he supposes that he has to do in the next trial in order to obtain another high score.

(b) By bringing into operation the 'directive effect', which has just been described. Judging from the introspective reports of the subjects this tendency also was sometimes clearly conscious, and involved formulation in words of the kind of modification of behaviour which would be required in the next trial; but, as the experimental conditions became more familiar, the tendency to make the appropriate correction became more 'automatic'.

(c) By setting up a general attitude of 'interest' or 'keenness'. This appears to be conducive to careful and accurate performance.

With regard to the effect of removal of knowledge of results, the most striking feature of the results is the surprisingly rapid deterioration of performance in Exp. I. In Exp. II the deterioration was less rapid, but it occurred with nearly all the subjects. These results seem to support the view that not only the acquisition of a skill, but also its maintenance, depends upon a continual 'check up' on the accuracy of the movement which has been made. It is further evidence, if further evidence is needed, against the classical view of a habit as something which will continue to go off automatically without any reference to the effects which it produces. Those subjects who succeeded in maintaining a high level of performance for any length of time after knowledge of results had been removed were very probably using, consciously, some clue, such as the kinaesthetic sensations from the wrists at the end of each movement, as a check on that movement, and as a guide to the nature of the correction required in the next movement.

The details of the performance of the subjects in these experiments raise several questions of general importance in experimental studies of learning. The task which the subjects were required to carry out was a relatively simple one; but, whether we regard their objective performance or their introspective reports, it is clear that they carried out this task by a surprising variety of methods. The differences of method were so marked that it would almost be fair to say that the different subjects were learning different tasks. Learning to move two handles separately so that an imaginary spot of light pursues a certain path can hardly be regarded as the same task as learning to apply a certain 'amount of force' simultaneously to the two handles—even though the end result which is aimed at is the same in both cases. This obviously raises the

questions of how far, in experimental studies of learning, it is useful to pool together the results of different subjects using different methods; and of how far it is possible either to impose a uniformity of method upon different subjects, or to obtain (by introspection or other means) a reliable indication of the nature of the differences in method between different subjects. In further experiments which are now in progress at the Cambridge Psychological Laboratory on the effect of knowledge of results, an attempt is being made to find whether greater objective simplicity in the task which the subject has to perform leads to greater uniformity in the methods adopted by different subjects.

VI. SUMMARY

Using an apparatus in which the subjects attempted, by a movement of the two hands, to direct a spot of light on to the bull's-eye of a target, a study has been made of the way in which knowledge of results affects the acquisition and maintenance of a muscular skill. (Knowledge of results could be removed by arranging that the subject was unable to see what part of the target he had struck.)

It was found (*a*) that no improvement in accuracy of performance occurs without knowledge of results, (*b*) that improvement occurs with knowledge of results, and (*c*) that removal of knowledge of results after skill has been acquired leads to a deterioration of performance.

Knowledge of results appears to lead to improvement in performance, (*a*) by causing a tendency to repeat actions which have been successful, (*b*) by what may be called a 'directive effect', i.e. by causing a tendency to correct, in the appropriate direction, any unsuccessful actions, and (*c*) by setting up a conscious attitude or mood which is conducive to accurate performance. Removal of knowledge of results produces, on the other hand, an attitude which is not conducive to accurate performance.

Certain implications of these results are discussed.

We wish to thank Prof. F. C. Bartlett, who suggested these experiments, for his help and advice.

REFERENCES

- ARPS, G. F. (1917). "Work with knowledge versus work without knowledge of results." *Psychol. Rev.* xxiv, 449.
- BLODGETT, H. C. (1929). "The effect of introduction of reward upon the maze performance of rats." *Univ. Calif. Publ. Psychol.* iv, No. 8.
- BRUCE, R. H. (1930). "The effect of removal of reward on the maze performance of rats." *Univ. Calif. Publ. Psychol.* iv, No. 13.
- CRAWLEY, S. L. (1926). "An experimental investigation of recovery from work." *Arch. Psychol.* xiii, No. 85.
- GRINDLEY, G. C. (1932). "The formation of a simple habit in guinea pigs." *Brit. J. Psychol.* xxiii, 127.
- JOHANSON, A. M. (1922). "The influence of incentives and punishment upon reaction time." *Arch. Psychol.* viii, No. 54.
- THORNDIKE, E. L. (1931). *Human Learning*. New York: Century.
- (1932). *The Fundamentals of Learning*. New York: Bureau of Publications, Teachers' College, Columbia University.
- (1935). *The Psychology of Wants, Interests, and Attitudes*. New York: Century.
- TOLMAN, E. C. (1932). *Purposive Behaviour in Animals and Men*, Chapter III. New York: Century.

(*Manuscript received 24 September 1937*)

SPEED OF WORK IN INTELLIGENCE TESTS

BY PATRICK SLATER

From the Department of Psychology, University College, London

- I. *Introduction* (pp. 55-56).
- II. *Materials and cases for the experiments* (pp. 56-57).
- III. *Methods of procedure* (pp. 58-61).
- IV. *Results and analysis* (pp. 61-68).
- V. *Summary and conclusions* (p. 68).
- References* (p. 68).

I. INTRODUCTION

THE following is a short account of some of the results obtained in a research into the relationship between speed, accuracy and level of ability in intelligence tests. The experimental studies were carried out on 450 children in Willesden County School during 1934-5, as part of the work for a Ph.D. thesis. They included studies of the relationships between speed and perseveration and fluency, as well as performances in intelligence tests, both verbal and non-verbal. But the present statement is confined to the results obtained from the measurements of the speed rates of several groups when working without fixed time limits, and of their intelligence both when time limits were fixed and when they were not. The shortness of this summary account unfortunately makes it impossible to discuss references in detail.

Much previous work on speed has assumed that speed rates are most clearly shown when subjects are encouraged to work as fast as possible, regardless of the quality of their work. If, however, intelligence is what underlies adaptability to different types of test situation, it may be supposed that given such encouragement the most intelligent subjects will work as fast as possible, whether they normally do so or not, and that only those who are backward in intelligence will work slowly. Instructions to work as fast as possible need not bring out the natural speed rates, but may obscure them. The natural speed rates will be most clearly revealed when subjects are not instructed to work specially quickly, but are allowed to choose their own rates.

The person who solves many problems in a test in an allotted time is

often assumed to have a higher speed rate for solving each particular problem than the person who has solved only a few. The total output is thus taken to have a close relation to the speed at which each particular item constituting the output is solved. But although at the end of a given time one individual may have furnished a larger number of correct solutions than another, it does not necessarily follow that he has worked faster. The one who has done few right may have worked quickly, but spent much of his time on incorrect or uncompleted work; the one who has done many right may have worked slowly, but made few mistakes. Thus speed rates cannot be measured by taking the total number of correct solutions furnished in a given time, whether that time be short or long.

To avoid this difficulty measurements of total numbers of solutions given, whether correct or not, have been taken as the basis for some experiments. This leads only to confusing incomparable measurements. Times spent on incorrect solutions not only cannot be compared with times spent on correct solutions; they have also no common basis for comparison with one another. Only correct solutions provide a common basis for the comparison of times.

For the measurement of speed rates in these experiments, two principles have been observed:

- (1) That only the times for correct solutions should be used.
- (2) That the subjects should be allowed to work at their own rates.

II. MATERIALS AND CASES FOR THE EXPERIMENTS

In order to measure both speed rate and level of ability, an adaptation of the Thorndike CAVD tests was prepared (3). This test consists of four types of problems: Completion, Arithmetic, Vocabulary and Directions. The difficulty of each problem has been measured separately from a theoretical zero. The problems are grouped in levels of difficulty, like those of the Binet-Simon test. But there are the following differences: In the CAVD tests the same types of problem, with a few exceptions, are used at all levels, and difficulty is intended to increase by equal intervals from each level to the next.

In the adaptation used five problems were selected from each level in the Completion series from level F to level Q inclusive, making a total of 60 problems. In the Arithmetic series, five were selected from each level from F to I inclusive, four from each level from J to N (excluding level L, which contains problems of a slightly different character from

those used in the other levels), three from level O and one from level P, making a total of 40 problems. In the Vocabulary series six questions were taken from each level from F to Q inclusive, making 72; but as the time taken in answering them is often very short, the questions for each level were combined in groups of three, making 24 groups altogether. In the Directions series four problems were taken from each level from J to Q inclusive, making 32 questions. Level Q is the highest in the original tests. The levels attained by the children, as judged from their total performances, varied between K and N.

The selection of the problems could not be altogether left to pure chance, as might have been theoretically desirable, since some of the questions referred to dollars, ice-cream, and other institutions more familiar to American children than to English.

In addition to the measurements of speed rate and of level of ability obtained from the CAVD tests (referred to as Completion S, Completion L, etc.), measurements of intelligence were obtained from various non-verbal tests, which are described here as Classification, Classification A, Code, Intersections A and Patterns. These were all done under fixed time limits. Intersections A is the only entirely new test. The types of problem on which the others were based have been used in accepted non-verbal tests, and their general usefulness as reference 'g' tests shown. The results of the two tests referred to as 'A' tests are scored on a special principle: Success is measured not from the total number of responses, but from the ratio of correct responses to incorrect ones. This method was originally adopted in order to measure accuracy of work, regardless of total output, but the measurements proved suitable for use among the reference 'g' tests. Classification A is the same test as Classification, but scored on this principle.

The children tested were those in the III and IV forms of the school. There are four IV forms, IV A, IV α , IV B and IV β , all nominally parallel; in each of them the children are of approximately the same age (about $14\frac{1}{2}$) and work on the same curriculum, but the more promising ones are put into the A and α forms. The four III forms, III A, III α , III B and III β , are similarly constituted, but the children are about a year younger. In all forms boys and girls are fairly evenly mixed. The total number of children in the IV forms was 106, in the III forms, 120.

III. METHODS OF PROCEDURE

In the CAVD tests the children were allowed to work at their own rate. An apparatus was installed to enable each child to record for himself the amount of time he spent on each problem. From the outside this apparatus, which was called the 'Number Clock', somewhat resembles a giant cyclometer. It consists of three cylinders, rotated by an universal electric motor. The outside rim of each cylinder is numbered 0, 1, 2, ..., 9, each number being 4 in. high (the same size as those used on the number-plates of cars). The motor is adjusted to make one complete revolution every 2 sec., and to make contact with the first of the three cylinders (reading from right to left) during a tenth part of this revolution. This moves the cylinder on one place and increases the unit digit by 1. When the digit on the unit cylinder changes from 9 to 0, the motion is conveyed on to the next cylinder. The three cylinders, taken together, record up to 999 units of time—the equivalent of 33 min. 20 sec. Each number remains motionless for 1·8 sec. and is changed in 0·2 sec., whether one digit only is involved, or all three. When the number is 999 the next revolution turns it back to 000, so that the machine can continue running indefinitely. The whole is enclosed in a box, with a slot in the side to leave one number on each of the cylinders visible.

In addition to the precautions always necessary in group testing, special measures have to be taken to ensure that everyone in the classroom can see the machine and read the time without difficulty. The machine is then switched on and its purpose explained. The first series of questions is given out, with a practice sheet affixed in front. On this there are examples of the kinds of questions used in each of the four series, with instructions on how to do them. The instructions about the times are simply: "Put down the time at which you start each question." Space for this is provided on the left of the page beside each question, thus:

Time

..... 1.	7 and 15 make	Answer
..... 2.	13 from 21 leaves	Answer

Verbal instructions, in so far as they affect the times, are: "For these questions you have as much time as you need. Read the instructions carefully before you begin. Then write down the time you start. When you have finished the first question, write down the time you start the next. Do not leave a question before you have finished it. Once you have gone on, you cannot go back." The children then do the examples for

the first test. The examiner checks up to see that everyone in the room has carried out the instructions correctly. Instructions may be repeated as often as necessary in order to ensure that everyone understands exactly what has to be done before the real test is started. When the last of the examples has been done, the children are again reminded of what they have to do, and the additional instruction is given: "When you have finished this test, hold up your hand for the next. Now turn over." The children thus start on one series and work through it until they can do no more before going on to the next. No verbal instructions are given after the first series has been begun. (The children who were tested in this way had already done questions of the CAVD types under other conditions.)

These instructions show that the time taken over a problem is measured as the difference between the time when it was begun and the time when the next was begun. Any pause a subject may make, after he has finished one problem and before he goes on to the next, becomes included in his total time. This might seem a source of possible error in the time measurements. But such pauses may be a functionally significant part of the whole reaction, and it is reasonable to extend the definition of "time taken" so as to include them. If the children were allowed to record separately the times they started and finished each problem, they would be exposed to a temptation to cheat; and whether this were proved or not, it would throw doubt on the whole procedure. The only effective alternative is to test each child separately and record his times by direct observation. But this is not possible in a school. Information about the nature of any special work spreads around far too quickly.

The children's rates of work were calculated separately for each of the four series—Completion, Arithmetic, etc. The procedure is as follows: All the times taken to solve a given problem correctly are summed, and the average time and standard deviation calculated. Let the average time = x , the standard deviation = σ_d , the time taken by any individual = $x + d$. Then the individual's speed rate for the problem is taken as

$$\frac{d}{\sigma_d}.$$

If he has failed to solve the problem correctly, no regard is paid to the time he spent on it. In the Vocabulary series, where the items are arranged in groups of three, only those times are considered in which all three items have been answered correctly.

Let the number of problems in any series which has been correctly

answered by any individual = t . Then t measures of his speed rate on that series will be obtained. These may be represented by the series:

$$\frac{d_1}{\sigma_{d_1}} + \frac{d_2}{\sigma_{d_2}} + \dots + \frac{d_t}{\sigma_{d_t}}.$$

If he has no constant rate of work, these values should tend to vary at random, some being positive and others negative, so that the sum of the series would tend to = 0. Actually it was found that in the groups tested the individuals tended to adopt regular rates of work, some faster, some slower than the average. Individuals who work quickly at easy problems also work quickly at hard ones, and those who are slow at the beginning continue slowly to the end.¹

If the subject works quickly the sum of the measurements of his speed rate is negative, if slowly, positive. His average speed rate can be expressed as:

$$\frac{1}{t} \sum \frac{d}{1-t\sigma_d}.$$

This average rate of work is calculated for each subject in each series (CAVD) in which he has successfully answered any considerable number of the problems, and constitutes his speed-rate score for that series. A 'high' speed rate, in the present context, means one with a negative value, and a 'low' speed rate one with a positive value. A positive correlation between a speed-rate measurement and a measurement of intelligence in the tables below indicates that the children whose speed rates are high tend to score highly in the intelligence test.

Since there are 156 possible speed-rate measures for each individual, this method of calculation is somewhat laborious. But it eliminates times spent on problems which are not correctly answered, and the differences in the average times and the variabilities of the times for different problems. Speed rates are measured with a degree of accuracy and reliability which does not seem to have been previously attained.

Scores in the level tests are based simply on the number of questions correctly answered. Difficult problems are given no greater weight than

¹ The speed rates in each series have been calculated from times spent on problems of every degree of difficulty from very easy to very hard. The correlation between speed rates on the easy and the difficult problems in any series is thus in the nature of a reliability coefficient for the series as a whole. But it is known from general theory that if two measurements correlate together, the reliability coefficient for each separate measurement cannot be significantly less than the correlation between the two. Since it is shown in the text that there is positive correlation between the speed-rate measurements obtained from different series, there is no need to advance separate proof that intercorrelation exists between the different members of the same series.

easy ones. For if success at a difficult problem is a sign of high intelligence, failure at an easy one must be a sign of low intelligence—one sign cannot be emphasized and the other neglected; and if failure at the easy problem is due to chance, so too may be success at the difficult one. The fact that subjects who succeed at some difficult problems fail at some easy ones may indicate either that their intelligence varies in relation to the test problems, or that the difficulty of the problems varies in relation to different subjects, or both. In assessing the general level of the subject's intelligence, such discrepant results should be balanced against one another, and their extent and significance studied as a separate problem. (This is done in another part of the present research.)

IV. RESULTS AND ANALYSIS

As the children had as much time as they needed to work at each of the CAVD tests, many did not finish all four. The results therefore cannot be assembled into two large groups, one for the III and one for the IV forms. Further, the material adapted from the Directions test proved not to have the same amount of difficulty for these subjects as it had for those on whom it had originally been standardized, and measures obtained from it had very little relation to those obtained from any of the other tests. Accordingly the Speed and Level measures for the following tables are taken from the results of the Completion, Arithmetic and Vocabulary tests only.

The results for the IV forms are in three groups:

IV forms CA group—49 cases, scored for speed rate and level of ability on the Completion and Arithmetic tests (Table V).

IV forms CV group—55 cases, scored on the Completion and Vocabulary tests (Table VI).

IV forms AV group—42 cases (Table VII).

These groups are not mutually exclusive. There are 37 cases common to all three. This must be remembered when the significance of the results is estimated. But as different tests are paired together each group provides some information not to be found in any of the others; and by comparing different groups general tendencies can be distinguished from chance fluctuations.

The results for the III forms are in two groups:

III forms CA group—70 cases (Table III).

III forms CV group—65 cases (Table IV).

Of these, 47 cases are common to both.

Speed of Work in Intelligence Tests

There is also a sixth group: III forms CA—82 cases. This includes all the 70 cases of Table III. The others are those of children who did not finish both the C and A tests. Their speed rates can be measured reliably from their incomplete results, but not their levels of ability. This table rests on the largest number of cases, is the shortest and the simplest to interpret, and is therefore given first. For the

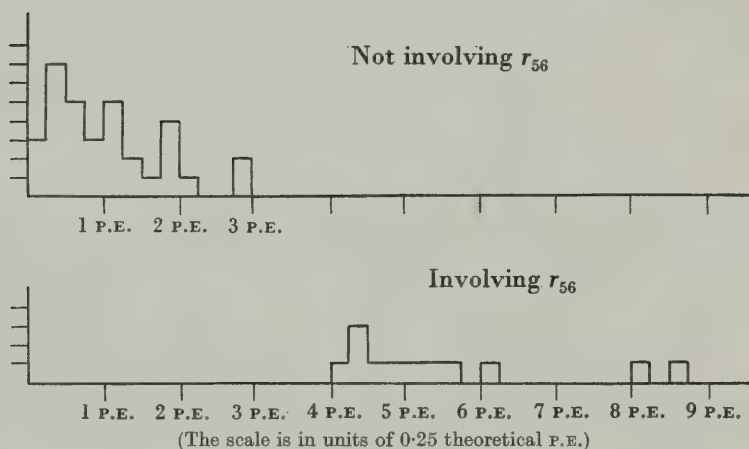
Table I. *III forms CA group, 82 cases*

<i>Correlations</i>					
	2	3	4	5	6
1. Code	0.596	0.318	0.374	0.246	0.137
2. Patterns	—	0.324	0.353	0.065	0.117
3. Classification	—	—	0.274	0.118	0.018
4. Intersections A	—	—	—	0.017	-0.090
5. Arithmetic S	—	—	—	—	0.516
6. Completion S	—	—	—	—	—

Probable errors of tetrad differences

Theoretical P.E. = 0.0345

	Observed P.E.	Obs. P.E. ÷ Theor. P.E.
1. All the tetrads	0.0745	2.160
2. Involving r_{56}	0.1345	3.898
3. (1) without (2)	0.0318	0.922

Fig. 1. *III forms CA group: tetrad differences.*

analysis Spearman's method is used, since its hypotheses are widely known and fairly generally accepted. Tetrad differences, systematically used, provide fuller material for locating the presence of disturbing factors than can many other methods of factor analysis; and in these tables, since only a few tests are involved, the number of tetrads is kept within moderate limits.

In Table I it is immediately clear that there is a special degree of correlation between the two measures for speed rate; and the distribution of the tetrad differences (see Fig. 1) shows that there is no need to suppose the presence of more than one general factor to account for the remaining correlations.

On these hypotheses, namely that:

- (1) the intercorrelations between tests 1-4 and the correlations between these tests and the speed-rate measures are due to one common factor, 'g', and that
 - (2) the correlation between the two measures of speed rate is influenced by another factor, common only to these two,
- the table is analysed, and the 'g' correlations and residual correlations shown in Table II obtained.

Table II. 'g' correlations and residuals derived from Table I

'g' correlations					
r_{1g}	0.925		r_{4g}	0.365	
r_{2g}	0.728		r_{5g}	0.163	
r_{3g}	0.471		r_{6g}	-0.041	
Residual correlations					
	2	3	4	5	6
1. Code	-0.078	-0.118	0.036	0.095	0.175
2. Patterns	—	-0.019	0.087	-0.054	0.147
3. Classification	—	—	0.102	0.041	0.038
4. Intersections A	—	—	—	-0.042	-0.057
5. Arithmetic S	—	—	—	—	0.523
6. Completion S	—	—	—	—	—

Note. If $r=0$, $\sigma_r=0.111$.

Tables III-VII show the same general tendencies as Table I, and do not need individual analysis.

In all these groups the measures of speed rate are significantly correlated together, but their crude correlations with the measurements from the non-verbal tests and the level tests are generally low. The tetrads into which the speed-rate correlations enter are large, and their differences are high. Individual tetrad differences larger than 5 times the theoretical P.E. occur in all the tables (except perhaps Table V).

In general, the C, A and V tests for level appear specially closely correlated. This is more conspicuous among the IV forms than among the III. It is not necessary to suppose the existence of any new special factor to explain this. The CAVD tests are verbal, and the existence of a 'v' factor has been shown. It is also known to be more evident among older groups. It would in fact have been surprising if no special correlation had been found here; as it has, the hypothesis that it is due to v factor seems the most reasonable one to adopt.

Table III. *III forms CA group, 70 cases*

	<i>Correlations</i>						
	2	3	4	5	6	7	8
1. Arithmetic L	0.272	0.251	0.353	0.414	0.370	—	0.024
2. Completion L	—	0.353	0.295	0.287	0.374	0.194	—
3. Classification	—	—	0.312	0.296	0.327	0.174	0.068
4. Intersections A	—	—	—	0.306	0.309	-0.088	-0.168
5. Patterns	—	—	—	—	0.585	0.138	0.219
6. Code	—	—	—	—	—	0.204	0.141
7. Arithmetic S	—	—	—	—	—	—	0.457
8. Completion S	—	—	—	—	—	—	—

Probable errors of tetrad differences

Theoretical P.E. = 0.0356

	Observed P.E.	Obs. P.E. ÷ Theor. P.E.
1. All the tetrads	0.0560	1.574
2. Those involving r_{78}	0.1092	3.068
3. (1) without (2)	0.0422	1.185

Table IV. *III forms CV group, 65 cases*

	<i>Correlations</i>						
	2	3	4	5	6	7	8
1. Completion L	0.452	0.323	0.140	0.255	0.143	—	0.159
2. Vocabulary L	—	0.027	0.346	0.210	0.218	0.031	—
3. Classification	—	—	0.398	0.399	0.335	-0.033	0.069
4. Intersections A	—	—	—	0.387	0.252	-0.069	0.006
5. Patterns	—	—	—	—	0.672	0.196	0.195
6. Code	—	—	—	—	—	0.214	0.259
7. Completion S	—	—	—	—	—	—	0.676
8. Vocabulary S	—	—	—	—	—	—	—

Probable errors of tetrad differences

Theoretical P.E. = 0.0383

	Observed P.E.	Obs. P.E. ÷ Theor. P.E.
1. All the tetrads	0.0792	2.068
2. Those involving r_{78}	0.1642	4.284
3. (1) without (2)	0.0544	1.421
4. Those involving r_{12}	0.0870	2.271
5. (3) without (4)	0.0462	1.206

Table V. *IV forms CA group, 49 cases*

	<i>Correlations</i>						
	2	3	4	5	6	7	8
1. Completion L	0.433	0.388	0.188	0.102	0.332	—	0.160
2. Arithmetic L	—	0.276	0.215	0.095	0.054	-0.165	—
3. Classification A	—	—	0.442	0.220	—	0.117	0.122
4. Patterns	—	—	—	0.355	0.294	0.063	0.154
5. Code	—	—	—	—	0.352	0.292	0.177
6. Classification	—	—	—	—	—	0.302	0.476
7. Completion S	—	—	—	—	—	—	0.458
8. Arithmetic S	—	—	—	—	—	—	—

Table V (*continued*). *Probable errors of tetrad differences*

	Theoretical P.E. = 0.0410	
	Observed P.E.	Obs. P.E. ÷ Theor. P.E.
1. All the tetrads	0.0604	1.471
2. Those involving r_{12}	0.0869	2.117
3. (1) without (2)	0.0543	1.323
4. Those involving r_{78}	0.0815	1.985
5. (3) without (4)	0.0470	1.144

Note. This is the least conclusive of the six tables: it shows the same general characteristics as the others, but not so clearly. The only tetrad larger than 5 times the theoretical P.E. is $r_{12}r_{78} - r_{18}r_{27}$. On the general rule followed in these tables, this tetrad has been assigned to the correlation which has the largest other tetrads—in this case, r_{12} .

Table VI. *IV forms CV group, 55 cases*

	<i>Correlations</i>						
	2	3	4	5	6	7	8
1. Completion L	0.637	0.295	0.087	0.225	0.369	—	0.315
2. Vocabulary L	—	0.146	0.039	0.151	0.359	0.278	—
3. Classification A	—	—	0.376	0.268	—	0.019	0.280
4. Patterns	—	—	—	0.489	0.371	0.026	0.077
5. Code	—	—	—	—	0.327	0.208	0.253
6. Classification	—	—	—	—	—	0.237	0.362
7. Completion S	—	—	—	—	—	—	0.519
8. Vocabulary S	—	—	—	—	—	—	—

Probable errors of tetrad differences

	Theoretical P.E. = 0.0409	
	Observed P.E.	Obs. P.E. ÷ Theor. P.E.
1. All the tetrads	0.0713	1.744
2. Those involving r_{12}	0.1200	2.936
3. (1) without (2)	0.0585	1.432
4. Those involving r_{78}	0.0968	2.369
5. (3) without (4)	0.0472	1.154

Table VII. *IV forms AV group, 42 cases*

	<i>Correlations</i>						
	2	3	4	5	6	7	8
1. Arithmetic L	0.510	0.490	0.207	0.219	0.299	—	0.029
2. Vocabulary L	—	0.138	0.072	0.046	0.241	0.162	—
3. Classification A	—	—	0.356	0.160	—	0.210	-0.021
4. Patterns	—	—	—	0.384	0.299	0.026	0.032
5. Code	—	—	—	—	0.163	-0.263	0.103
6. Classification	—	—	—	—	—	0.252	0.325
7. Arithmetic S	—	—	—	—	—	—	0.569
8. Vocabulary S	—	—	—	—	—	—	—

Probable errors of tetrad differences

	Theoretical P.E. = 0.0448	
	Observed P.E.	Obs. P.E. ÷ Theor. P.E.
1. All the tetrads	0.0695	1.552
2. Those involving r_{78}	0.1195	2.666
3. (1) without (2)	0.0562	1.254
4. Those involving r_{12}	0.0827	1.846
5. (3) without (4)	0.0492	1.097

When the tetrad differences involving correlations between level measures or speed-rate measures have been isolated, the remainder are none as great as 5 times the theoretical P.E. But their observed P.E. is still slightly larger than the theoretical P.E. In all the tables, some correlations have had to be omitted (e.g. Completion L—Completion S), since they are liable to special disturbances of no general significance. The total number of tetrad differences is thus reduced from 210 to 152 in Tables III and IV, from 210 to 126 in Tables V, VI and VII. This tends to increase the size of the observed P.E.'s. In Table III, the remaining tetrads, after those involving the speed-rate correlation have been isolated, have an observed P.E. 1.185 times as large as the theoretical P.E. But since all the cases in this table are also in Table I, where there are no omitted values, and where the corresponding P.E. is smaller than the theoretical P.E., it seems reasonable to treat the ratio 1.185 as within the range of chance variation.¹ If this is so the corresponding ratios in the other four tables may also be treated as chance variations, since all are closely comparable (Table IV, 1.206; Table V, 1.144; Table VI, 1.154; Table VII, 1.097). Thus no other correlations save those involving the level tests or the speed-rate measures need be supposed to produce significant disturbances among the tetrad differences.

This may be verified in another way. To reduce the ratio of observed P.E. to theoretical P.E. further, the tetrads to isolate in Table III would be those from r_{48} (Intersections A—Completion S); but if they were generally significant, the tetrads from r_{47} in Table IV would also need isolation. This is not the case. In Table IV, on the other hand, the tetrads to isolate would be those from r_{56} ; but this is not necessary in Table III. In short, the lesser disturbances show no consistency; they may properly be regarded as due to chance.

The same general hypotheses are therefore adequate for the interpretation of all the tables:

(1) That a special factor (probably 'v') is present among the C, A and V level measures.

(2) That there is a factor of speed rate, or speed preference, among the speed-rate measures.

(3) That the intercorrelations between the level and the speed-rate measures, the correlations of both with the non-verbal tests, and of the non-verbal tests among themselves are due to one common factor, 'g'.

¹ This part of the analysis is hampered by lack of appropriate formulae for computing P.E.'s of P.E.'s.

On these hypotheses, the 'g' correlations and residuals shown in Table VIII are obtained.

Table VIII. 'g' correlations and residuals of the level and speed-rate measures given in the preceding tables

	'g' correlations					
Table ...	I	III	IV	V	VI	VII
Completion L	—	0.563	0.368	0.388	0.436	—
Arithmetic L	—	0.552	—	0.096	—	0.481
Vocabulary L	—	—	0.295	—	0.327	0.247
Completion S	-0.033	-0.054	0.061	0.131	0.282	—
Arithmetic S	0.141	0.191	—	0.374	—	-0.203
Vocabulary S	—	—	0.228	—	0.438	0.149

	Speed rate			'Verbality'		
	CA	CV	AV	CA	CV	AV
Table I	0.444	—	—	—	—	—
Table III	0.467	—	—	(-0.039)	—	—
Table IV	—	0.662	—	—	0.344	—
Table V	0.409	—	—	0.401	—	—
Table VI	—	0.395	—	—	0.495	—
Table VII	—	—	0.599	—	—	0.392

The residual correlations of the speed-rate measures indicate that the 'S' correlation of Completion S is about 0.60, of Arithmetic S about 0.70, and of Vocabulary S about 0.85.

'Classification' can be compared with 'Classification A' in Tables V, VI and VII. 'Classification A' generally correlates higher with the measures of intelligence and lower with the measures of speed rate than 'Classification'. A possible explanation is that 'Classification' measures are influenced by speed rate. The tetrad differences lend some support to this view—particularly those involving r_{68} in Table V. The evidence, however, is not strong enough to support any definite or general conclusions.

Although it shows closer connexions with speed rate than the other tests, 'Classification' did not demand quicker work. The time for doing it was limited, but the limits were generous, and children were not pressed to work fast. The test calling for greatest quickness of judgement, under pressure, was the Code test. Yet this test does not show specially high correlations with the speed-rate measures, and proves to be the best of the reference 'g' tests. In this respect the results of the present experiments support the findings of Bernstein⁽¹⁾ and Hunsicker⁽²⁾: they provide no evidence that conditions of pressure vitiate intelligence tests by bringing a special speed factor into play. There is also no evidence that the factor common to the speed-rate measures used here is associated with conditions of pressure.

As has already been mentioned, the four III forms are all nominally parallel, and so are those of the IV; but in practice the more promising children were put into the A and α forms and the less promising into the B and β . In particular, form IV β was composed of a small number of noticeably backward children, whose education required special attention. The test results show that the 10 children in this form who did the Completion test have an average level only $0.988\sigma^1$ below the 78 children in the other IV forms who did this test, but a speed rate 3.498σ below the others. Similarly, the 11 children who completed the Arithmetic test had an average level only 0.026σ below the 51 others who did it, but their speed rate was 2.484σ lower. These facts seem to suggest that exceptionally low speed rates may at times be a more serious source of difficulty to the educationalist than low levels of ability.

V. SUMMARY AND CONCLUSIONS

A method is described for measuring speed rates at which children work when doing problems in intelligence tests. This method enables them to work at their own pace and to record for themselves the time they spend on each problem. Measures of speed rate obtained by this procedure from various groups agreed among themselves, but did not show any close association with measures of intelligence obtained either from verbal or non-verbal intelligence tests, whether the time for doing them was limited or not.

These statistical facts are evidence that among the groups studied the children, when left free to set their own pace of work, tended to adopt consistent speed rates. They adhered to these when working on problems of different degrees of difficulty and of different types. Their speed rates could not be considered to depend only on the amount of their general intelligence.

Some evidence was obtained which suggests that the study of speed rates may be of value for educational psychology.

REFERENCES

- (1) BERNSTEIN, E. (1924). "Quickness and intelligence." *Brit. J. Psychol. Monog. Suppl.* III, No. 1.
- (2) HUNSICKER, L. M. (1925). *A Study of the Relationship between Rate and Ability.* Teachers' College, Columbia University.
- (3) THORNDIKE, E. L. and others (1926). *The Measurement of Intelligence.* Teachers' College, Columbia University.

¹ This is the σ of the difference between the means.

(Manuscript received 15 July 1937)

NOTE ON PROFESSOR GODFREY H. THOMSON'S ARTICLE "THE INFLUENCE OF UNIVARIATE SELECTION ON FACTORIAL ANALYSIS OF ABILITY"

BY WALTER LEDERMANN
Moray House, University of Edinburgh

- I. *A change in notation bringing all Thomson's results under one formula (pp. 69-71).*
- II. *A matrix proof of Thomson's theorem that the rank is unchanged (pp. 71-72).*
- III. *A formula for the new loadings (pp. 72-73).*
References (p. 73).

I. A CHANGE IN NOTATION BRINGING ALL THOMSON'S RESULTS UNDER ONE FORMULA

IN his paper⁽¹⁾ Prof. Godfrey H. Thomson discusses the changes in the intercorrelations of t tests, when the population to which the tests are administered is directly selected with respect to one of the tests, x_1 , in such a way that the standard deviation Σ_1 of that test is reduced to

$$\sigma_1 = p_1 \Sigma_1, \quad (0 < p_1 < 1).$$

In the parent population the analysis of the t tests is based on a symmetrical matrix

$$R_0 = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1t} \\ R_{21} & R_{22} & \dots & R_{2t} \\ \dots & \dots & \dots & \dots \\ R_{t1} & R_{t2} & \dots & R_{tt} \end{bmatrix}, \quad (R_{ij} = R_{ji}), \quad \dots\dots(1)$$

in which R_{ij} ($i \neq j$) is the correlation of test i with test j in standard measure. The diagonal cells R_{ii} are either all equal to unity ($R_{ii} = 1$) or equal to the communalities ($R_{ii} = H_i^2$; see (1), p. 456), according as we are considering the 'complete' matrix R_1 or the 'reduced' matrix R . Our notation will enable us to deal with both cases simultaneously.

70 *Note on Professor Godfrey H. Thomson's Article*

In the selected population, the matrix (1) is changed into

$$R_0^* = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1t} \\ r_{21} & r_{22} & \dots & r_{2t} \\ \dots & \dots & \dots & \dots \\ r_{t1} & r_{t2} & \dots & r_{tt} \end{bmatrix}, \quad (r_{ij} = r_{ji}), \quad \dots\dots(2)$$

where r_{ij} ($i \neq j$) denotes the intercorrelation of test i and test j , again in standard measure. By applying the general theory of selection Prof. Godfrey Thomson deduced the following values for the new correlations:

$$r_{1i} = \frac{p_1}{p_i} R_{1i}, \quad (i = 2, 3, \dots, t), \quad \dots\dots(3)$$

$$r_{ij} = \frac{R_{ij} - q_i q_j}{p_i p_j}, \quad (i, j = 2, 3, \dots, t; i \neq j), \quad \dots\dots(4)$$

where $q_1 = \sqrt{(1 - p_1^2)}$, $q_i = q_1 R_{1i}$; $p_i = \sqrt{(1 - q_i^2)}$

(see (1), equations (2), (3), (1)). By a slight change of notation we shall combine (3) and (4) in one formula. Evidently, the process of selection depends only on one parameter, e.g. the ratio σ_1/Σ_1 or the 'coefficient of shrinkage' $q_1 = (1 - \sigma_1^2/\Sigma_1^2)^{\frac{1}{2}}$. Using the latter magnitude as independent parameter we put

$$\left. \begin{aligned} q_1 &= q; & p_1 &= \frac{1 - q^2 R_{11}}{\sqrt{(1 - q^2)}}; \\ q_i &= q R_{1i}; & p_i &= \sqrt{(1 - q_i^2)} = \sqrt{(1 - q^2 R_{1i}^2)}, \quad (i > 1). \end{aligned} \right\} \quad \dots\dots(5)$$

Thus all quantities with subscripts greater than unity are unaltered, and in the case $R_{11} = 1$ our notation coincides with that of Prof. Godfrey Thomson. Our p_1 is the only quantity which differs from his, and then only when $R_{11} \neq 1$. For *his* p_1 we shall in what follows write $\sqrt{(1 - q^2)}$ to avoid confusion.

$$\text{Now let} \quad r_{ij} = r_{ji} = \frac{R_{ij} - q^2 R_{1i} R_{1j}}{p_i p_j}. \quad \dots\dots(6)$$

By (5) this is obviously identical with (4), when $i \neq j$ and $i > 1, j > 1$. Next suppose that $j = 1, i > 1$. We then get

$$r_{i1} = r_{1i} = \frac{R_{1i} - q^2 R_{11} R_{1i}}{p_1 p_i} = \frac{R_{1i}}{p_i} \frac{1 - q^2 R_{11}}{p_1},$$

$$\text{or by (5)} \quad r_{1i} = \frac{R_{1i}}{p_i} \sqrt{(1 - q^2)},$$

which is identical with (3). Thus (6) is equivalent to (3) and (4).

We shall now show that (6) holds also for the diagonal cells of the matrix R_0^* . When we are dealing with the complete correlational matrix, the diagonal entries both of R_0 and R_0^* must be unity, since we express all quantities in standard measure. Indeed, on putting

$$R_{11} = R_{22} = \dots = R_{tt} = 1$$

in (5) and (6), we find

$$r_{11} = r_{22} = \dots = r_{tt} = 1,$$

as can easily be verified. In the case of a reduced correlational matrix the elements $r_{ii} = h_i^2$ are determined by the condition that the rank of R_0^* shall be as low as possible. In his paper Prof. Godfrey Thomson shows that¹

$$h_1^2 = \frac{H_1^2 (1 - q^2)}{1 - q_1^2 H_1^2},$$

$$h_i^2 = H_i^2 - \frac{q_i^2}{p_i^2} (1 - H_i^2) = \frac{H_i^2 - q_i^2}{p_i^2}, \quad (i > 1)$$

(see (1), equations (4) and (5)). On writing H_i^2 for R_{ii} , and h_i^2 for r_{ii} in (5) and (6) we get

$$\left. \begin{aligned} h_1^2 = r_{11} &= \frac{R_{11} - q^2 R_{11}}{p_1^2} = (H_1^2 - q^2 H_1^4) \frac{1 - q^2}{(1 - q_1^2 H_1^2)^2} = \frac{H_1^2 (1 - q^2)}{1 - q_1^2 H_1^2}, \\ h_i^2 = r_{ii} &= \frac{R_{ii} - q^2 R_{1i}^2}{p_i^2} = \frac{H_i^2 - q_i^2}{p_i^2}, \end{aligned} \right\} \dots\dots (6')$$

in conformity with the above values. Hence (6) holds for all elements of R_0^* , including the diagonal elements.

II. A MATRIX PROOF OF THOMSON'S THEOREM THAT THE RANK IS UNCHANGED

For theoretical purposes it is convenient to express (6) as a matrix equation. We propose to construct a square matrix Z of order t such that

$$R_0^* = Z R_0 Z'. \quad \dots\dots (7)$$

The form of the right-hand member of (6) suggests that Z will be of the shape

$$Z = \begin{bmatrix} \frac{1+z_1}{p_1} & 0 & \dots & 0 \\ \frac{z_2}{p_2} & \frac{1}{p_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \frac{z_t}{p_t} & 0 & \dots & \frac{1}{p_t} \end{bmatrix}. \quad \dots\dots (8)$$

¹ We replace his p_1^2 by $1 - q^2$.

72 Note on Professor Godfrey H. Thomson's Article

On performing the triple matrix product in (7) and using (6) we obtain the conditions

$$\frac{1}{p_i p_j} (R_{ij} - q^2 R_{1i} R_{1j}) = \frac{1}{p_i p_j} (R_{ij} + z_i R_{1j} + z_j R_{1i} + z_i z_j R_{11}),$$

or
$$z_i z_j R_{11} + z_i R_{1j} + z_j R_{1i} + q^2 R_{1i} R_{1j} = 0.$$

To solve these equations we put

$$z_i = z R_{1i}, \quad (i = 1, 2, \dots, t), \quad \dots\dots(9),$$

where z is a quantity yet to be determined. On substituting for z_i and cancelling the factor $R_{1i} R_{1j}$ we find

$$z^2 R_{11} + 2z + q^2 = 0.$$

Hence
$$z = \frac{-1 + \sqrt{(1 - q^2 R_{11})}}{R_{11}}. \quad \dots\dots(10)$$

We have rejected the negative sign of the square root for this reason: in the limiting case when no selection takes place, we have $q = 0$, and the matrix Z should become the unit matrix, i.e. $z = 0$. This value is actually obtained from (10) by choosing the positive sign of the square root, which we shall accordingly retain also for the general case. It will be observed that z is always real and negative. Thus the matrix Z is completely determined by (8), (9) and (10), and the relation (7) is established.

The determinant of Z is readily calculated; we find

$$\det |Z| = (p_1 p_2 \dots p_t)^{-1} (1 + z_1),$$

or by (9) and (10)

$$\det |Z| = (p_1 p_2 \dots p_t)^{-1} \sqrt{(1 - q^2 R_{11})}.$$

This value is non-zero, i.e. Z is a non-singular matrix. Since pre- and post-multiplication by a non-singular matrix leaves the rank of a matrix invariant (see, e.g. Bôcher⁽²⁾, p. 29), we infer from (7): *The minimum rank of a correlational matrix is unaltered by selection, if the communalities are defined by (6')*. This is Prof. Godfrey Thomson's theorem (1, p. 453).

III. A FORMULA FOR THE NEW LOADINGS

We shall now apply the results of the preceding section to the factor analysis of the correlational matrix. Suppose that R_0 is of rank r and has been resolved into factors, thus

$$R_0 = FF', \quad \dots\dots(11)$$

where

$$F = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1r} \\ A_{21} & A_{22} & \dots & A_{2r} \\ \dots & \dots & \dots & \dots \\ A_{t1} & A_{t2} & \dots & A_{tr} \end{bmatrix}$$

is the $t \times r$ matrix of loadings or saturations. Then by (7) we obtain

$$R_0^* = ZFF'Z' = (ZF)(ZF)',$$

$$\text{or} \quad R_0^* = F^*F^{*'}, \quad \text{.....(12)}$$

$$\text{where} \quad F^* = [a_{ip}] = ZF. \quad \text{.....(13)}$$

Now, (12) is of the same form as (11), whence we get the result: *if F is a matrix of loadings for the parent population, then a (possible) matrix of loadings for the selected population is obtained by pre-multiplying F by the matrix Z .*

If the set of loadings found in this way should turn out to be unsatisfactory from the psychological point of view, a rotation of the reference axes becomes necessary; the general solution is therefore of the form

$$F^* = ZFP,$$

where P is an arbitrary orthogonal matrix of order r . Resolving (13) into components we can write

$$a_{ip} = \frac{1}{p_i} (A_{ip} + z_i A_{ip}) \quad \left(\begin{matrix} i=1, 2, \dots, t \\ p=1, 2, \dots, r \end{matrix} \right). \quad \text{.....(14)}$$

In conclusion, we shall illustrate this method by applying it to one of the examples given in Prof. Godfrey Thomson's paper⁽¹⁾, p. 456). There we have $t=5$, $r=1$, and from the data we find p_i , z_i and a_{i1} by (5), (9) and (14):

i	1	2	3	4	5
A_{i1}	0.900	0.800	0.700	0.600	0.500
z_i	-0.306	-0.272	-0.238	-0.204	-0.170
p_i	0.803	0.817	0.864	0.902	0.933
a_{i1}	0.778	0.679	0.562	0.461	0.372

This is in agreement with the values obtained by directly factorizing the matrix R_0^* (1, p. 457), within the limits of accuracy of the computation.

REFERENCES

- (1) THOMSON, GODFREY H. (1938). "The influence of univariate selection on factorial analysis of ability." *This Journal*, xxviii, 451-9.
- (2) BÔCHER, M. (1907). *Introduction to Higher Algebra*. New York: The Macmillan Company.

(Manuscript received 22 August 1937)

INTELLIGENCE, CONCRETE AND ABSTRACT:¹ NOTE

BY W. P. ALEXANDER

ON p. 165 of the Monograph on "Intelligence, Concrete and Abstract" a table of tentative norms for the Performance Scale described in Appendix 2 is given for ages 11-19 years. It appears from subsequent data that the norms for ages 11 and 12 should be corrected, and it is desired to extend the norms below the age of 11 years.

I therefore give below a revised table of norms from 7 to 19 years. These are still to be regarded essentially as tentative norms. In the upper range from 13 to 19 years they have proved satisfactory. The Scale was designed to operate at ages 14-16, and would appear to have less reliability with younger children. Even so, repeated requests have indicated a need for an extension of the table of norms, and it is hoped the table now given may be helpful.

Table I. *Monthly age norms of performance for the scale described in Appendix 2 of "Intelligence, Concrete and Abstract"*

Boys' norms, roman type; girls' norms, italic.

Years	Months										
	0	1	2	3	4	5	6	7	8	9	10
7	58	59	60	61	62	63	64	65	66	67	68
	<i>51</i>	<i>52</i>	<i>53</i>	<i>54</i>	<i>55</i>	<i>56</i>	<i>57</i>	<i>58</i>	<i>59</i>	<i>60</i>	<i>61</i>
	70	71	72	73	74	75	76	77	78	79	80
8	<i>63</i>	<i>64</i>	<i>65</i>	<i>66</i>	<i>67</i>	<i>68</i>	<i>69</i>	<i>70</i>	<i>71</i>	<i>72</i>	<i>73</i>
	82	83	84	85	86	88	89	90	91	92	94
9	<i>75</i>	<i>76</i>	<i>77</i>	<i>78</i>	<i>79</i>	<i>80</i>	<i>81</i>	<i>82</i>	<i>83</i>	<i>84</i>	<i>85</i>
	96	97	99	100	101	103	104	105	107	108	109
10	<i>87</i>	<i>88</i>	<i>89</i>	<i>90</i>	<i>91</i>	<i>92</i>	<i>93</i>	<i>95</i>	<i>96</i>	<i>97</i>	<i>98</i>
	112	113	115	116	118	119	121	122	124	125	127
11	<i>100</i>	<i>101</i>	<i>102</i>	<i>103</i>	<i>105</i>	<i>106</i>	<i>107</i>	<i>108</i>	<i>109</i>	<i>110</i>	<i>112</i>
	130	131	133	134	136	137	139	141	143	145	147
12	<i>114</i>	<i>115</i>	<i>116</i>	<i>118</i>	<i>119</i>	<i>120</i>	<i>122</i>	<i>123</i>	<i>124</i>	<i>126</i>	<i>127</i>
	150	152	154	156	158	160	163	165	167	169	171
13	<i>130</i>	<i>132</i>	<i>134</i>	<i>135</i>	<i>137</i>	<i>139</i>	<i>140</i>	<i>142</i>	<i>144</i>	<i>145</i>	<i>147</i>
	175	177	179	181	183	185	188	190	192	194	196
14	<i>150</i>	<i>152</i>	<i>154</i>	<i>155</i>	<i>157</i>	<i>159</i>	<i>160</i>	<i>162</i>	<i>164</i>	<i>165</i>	<i>167</i>
	200	202	204	205	207	209	210	212	214	215	217
15	<i>170</i>	<i>172</i>	<i>174</i>	<i>175</i>	<i>177</i>	<i>179</i>	<i>180</i>	<i>182</i>	<i>184</i>	<i>185</i>	<i>187</i>
	220	222	224	225	227	229	230	232	234	235	237
16	<i>190</i>	<i>192</i>	<i>194</i>	<i>195</i>	<i>197</i>	<i>199</i>	<i>200</i>	<i>202</i>	<i>204</i>	<i>205</i>	<i>207</i>
	240	242	244	245	247	249	250	252	254	255	257
17	<i>210</i>	<i>212</i>	<i>214</i>	<i>215</i>	<i>217</i>	<i>219</i>	<i>220</i>	<i>222</i>	<i>224</i>	<i>225</i>	<i>227</i>
	260	262	264	265	267	269	270	272	274	275	277
18	<i>230</i>	<i>232</i>	<i>234</i>	<i>235</i>	<i>237</i>	<i>239</i>	<i>240</i>	<i>242</i>	<i>244</i>	<i>245</i>	<i>247</i>
	280	281	282	284	285	286	288	289	290	292	293
19	<i>250</i>	<i>252</i>	<i>254</i>	<i>255</i>	<i>257</i>	<i>259</i>	<i>260</i>	<i>262</i>	<i>264</i>	<i>265</i>	<i>267</i>

¹ "Intelligence, Concrete and Abstract." *Brit. J. Psychol. Monog. Suppl.* (1935), No. 19.

(Manuscript received 13 December 1937)

CRITICAL NOTICE

BY O. L. ZANGWILL

Psychology down the Ages. By C. SPEARMAN. 2 volumes. London: Macmillan & Co. 1937. Vol. I, pp. xi+454; Vol. II, pp. vii+355. 30s. net.

Prof. Spearman's recent and distinguished work represents a new departure in historical psychology. Here is no ordinary endeavour to "follow up the meanderings of doctrine as governed by historical and extrinsic influences, whether cultural, theological, personal or otherwise". Prof. Spearman has set himself the exacting task of portraying the theoretical development of psychology from Aristotle to the present day. His courageous mission is to depict "not the waves but the whole tide; not the course of psychology, but its development".

In Part A (*What Psychology is About*), Prof. Spearman surveys in masterly fashion the gradual evolution of psychological method. Here, admirably told, is the story of that protracted endeavour to elevate "common sense about the mind" to the status of a general science. From "cogitation over the mystery of life" we are led through primitive animism and the doctrine of the 'soul' to the beginnings of an independent analytical psychology. But this independence has not been achieved without renunciation, and two major 'renunciations' characterize the emergence of an independent psychology. In the first place, Prof. Spearman urges, the ultimate disparity between bodily and mental activity precludes the possibility of any final explanation of mind in terms of body. Psychology may reasonably hope to obtain aid from physiology, but "all the more ambitious attempts to rescue the science of mind by appeal to that of the body have ended only in disaster". Progress in psychology consequently presupposes the renunciation of physiology and the development of its own explanatory categories. In the second place, Prof. Spearman points out, the progress of psychological discussion has been constantly obscured and arrested by the intrusion of seemingly insoluble philosophical issues. In order to progress, psychology must emancipate itself from philosophy, but this emancipation, he stipulates, can never become complete. The "seemingly insoluble philosophical issues" are themselves inextricably bound up with the nature of the psychologist's own problems. "Completely deprived of them", he observes, "psychology would become ridiculous." In the face of this dilemma, the author appears to advocate a policy of opportunism: "Psychological science must be ready to include in its scope *whatever it can*."

In the fourth chapter, Prof. Spearman considers at some length the data and methods of contemporary psychology. Of particular importance is the author's explicit statement of his position in regard to the confused and controversial question of psychological analysis. Prof. Spearman distinguishes at the outset between 'real' and 'ideal' analysis. The latter, the 'analysis by intention' of Lewes, he accepts as a thoroughly justifiable procedure, provided that it be followed by synthesis. "Psychology", he writes, "...begins with such convenient portions of experience as a sensory percept or an emotion..."

and from these works both *downwards and upwards*." Intentional analysis gives rise to the primary material of psychology, which has to be sorted out, selected, further analysed, and put together again. Such analysis, he holds, is constantly being performed by the ordinary individual, and is by no means exclusively confined to the laboratory. It is interesting to observe that Prof. Spearman's position in regard to the problem of analysis is virtually identical with the standpoint of Bühler. The latter has likewise pointed out that the intentional abstraction of form and colour involves no change whatsoever in direct phenomenal experience. Abstractive analysis should therefore be sharply distinguished from the existential analysis so abhorrent to the *Gestalt* school.

The second part of *Psychology down the Ages* constitutes the most valuable historical section of the entire book. Here the author subjects to penetrating scrutiny the familiar concepts of psychology's stock-in-trade. There are wise and balanced chapters on 'classical intellect' and 'modern intelligence', a brilliant logical (and etymological) analysis of the concept of 'attention', and a number of chapters concerned with various aspects of the doctrine of faculties. In the concluding chapter (*Faculties or Chaos?*), Prof. Spearman sounds a severe and timely warning to his less rigorous contemporaries. The assumption of functional unity, he points out, has been accepted unquestioningly by the greater number of those who openly disavow the doctrine of faculties: "... we moderns", he writes, "are continually using language and doing deeds whose sole possible excuse lies in some assumed functional going together. Drop this, and the scope of all knowledge is reduced to that of particular cases. Generalities disappear; and therewith science. The whole of the psychology fabricated after this fashion comes clattering down like a house built of cards."

Part C (*How the Psyche is Constituted*) may perhaps be accepted as the definitive statement of Prof. Spearman's position in regard to the traditional problems of experimental psychology. This section comprises four chapters on perception, an analysis of thinking, a vigorous essay on mental dynamics, and numerous chapters bearing upon outstanding theoretical issues. In Chapter XII the author presents the view that perception is elaborated out of a simple affective consciousness (the 'primitive pathic state' of Cattell) under the two-fold influence of sensory stimulation and directed cognition. Any investigation of perception must, he believes, involve a certain measure of analysis, but Prof. Spearman vigorously rebuts the charge of 'sensory atomism' brought so frequently and indiscriminately against all who deny the priority of 'wholes'. In the first place, he points out, the comparison between atoms, which are substantival, and sense-qualities, which are attributive, is an entirely false analogy. Moreover, he maintains, the term 'sum' used in reference to perception can only mean the coexistence of discrete sensations, possibly in the same consciousness, but otherwise in complete disconnexion. Now not even Hartley, Prof. Spearman reminds us, upheld a 'summation' theory of sense perception in so crude and extreme a form. For him, as for the great majority of English empiricists, perception was to be envisaged as a highly complex integration of sensations and associations. "The analysis of *primitive* perception into 'sensations'", he concludes, "would seem to be a not unreasonable conception, if taken with its proper limitations."

In his treatment of relations, Prof. Spearman stands closest to Meinong and his Austrian followers. Deriving from Sully, he defines the perception of form as the direct apprehension of particular fundamentals in some specific inter-

relationship. Prof. Spearman makes his position clear by means of a simple illustration: A bi-dimensional pattern, such as is produced by two straight lines intersecting at right angles, may be apprehended either as a cross, i.e. as a form, or else as two lines crossing each other, that is to say, as a relation. Yet in both these cases, he maintains, *the fact perceived remains identical*. "Relations and forms are really two terms denoting the very same objective facts, but indicating more or less different manners of perceiving them." It is perhaps to be regretted that Prof. Spearman makes no mention of the numerous examples of form perception in which the perceptual content demonstrably *does* change with alteration in the manner of observation. For example, one may recall the 'flicker-fusion' experiment of Hartmann in which change in the mode of observation was found to result not only in a distinct alteration in phenomenal appearance, but also in a shift in the critical fusion period.

The familiar reversals of 'figure' and 'ground' are likewise interpreted by Prof. Spearman in terms of variation in the manner of observation. He rightly points out that a spontaneous reversal of this character must of necessity entail some alteration in the clearness and intensity of our knowledge of the relations between the principal structural aspects of the perceived pattern. It is not, however, altogether clear from his presentation whether this change in our knowledge is to be envisaged as a cause or as an effect of the perceptual reorganization. A similar criticism might also be raised against Prof. Spearman's treatment of spontaneous grouping in the visual field. Following Schumann, he rejects the notion of configurational determinants, and ascribes the phenomenon to the acquisition and operation of specific perceptual habits. "We might thus surmise", he writes, "that . . . spontaneous grouping . . . is in good conformity with probable situations of ordinary life."

Prof. Spearman's general theory of visual perception must of necessity lead him into conflict with the protagonists of the *Gestalttheorie*, and in Chapter xxiv he launches a vigorous attack upon 'The Confusion that is Gestalt Psychology'. In this chapter, he subjects some of the fundamental *Gestalt* categories to a much-needed and extremely penetrating examination. Even the most loyal supporters of this movement would have to confess to a certain amount of theoretical confusion involved in contemporary *Gestalt* psychology, and they should pay careful attention to Prof. Spearman's trenchant criticism. He raises a justifiable protest against the variety of meanings commonly attributed to the term *Gestalt* itself, and calls special attention to the equivocality of the Law of *Prägnanz*. Prof. Spearman concludes his chapter with the query that "...if Associationism may be regarded as Psychological Enemy No. 1, cannot Gestaltism put in a claim to be at any rate No. 2?!" One may venture to suggest that even the most hardened Public Enemies have sometimes been known to reform!

The first volume of *Psychology down the Ages* concludes on a serious, even pessimistic, note. If William James is right, and psychology nothing but the hope of a science, it follows that a history of psychology can scarcely aspire to be more than an essay in optimism. Prof. Spearman ruefully admits that his historical account has been on the whole less optimistic than most. "The belief in continual progress", he writes, "seems to rest largely on an illusion. The course of mental science has to a great extent been one of discovering, forgetting, and rediscovering. The amiable historians overlook the forgetfulness, and so arrive at unbroken discovery. They are as one who should suppose that, because each wave is moving forward, the tide must be doing so."

With the commencement of Volume II the outlook immediately begins to brighten. Following a preliminary disquisition upon the nature of scientific law, Prof. Spearman presents in Part D the outcome of his own outstanding attempts to introduce scientific law into general psychology. The laws of retentivity, inertia, control, constant output, fatigue and noëgenesis are expounded in full, together with admirable summaries of the experimental researches upon which these principles are presumably based. Prof. Spearman's general psychology, with which readers of his earlier works are already familiar, is here restated with great vigour, freshness and persuasion. Of particular interest are the two chapters on retentivity, although one may perhaps regret that there is so little attempt to relate the theory of 'dispositions' to kindred contemporary formulations, such as the 'schema' hypothesis of Bartlett or the 'dynamic trace theory' of Koffka. There is also a very minor error of interpretation on p. 53, where 'phenomenal regression' is attributed to the influence of assimilation. The trend of the experimental work leads one to believe that constancy does not depend directly upon the operation of past experience.

Part E (*What goes with What*) consists in a general survey of the entire field of individual differences. A preliminary critique of 'The New Typology' leads to the somewhat disheartening conclusion that "...these 'types'—as previously, the faculties, temperaments, and so forth—have failed to supply just what constitutes an irreducible minimum for science in general—namely, a proven account of *what goes with what*". Prof. Spearman lays particular emphasis upon the critical need of individual psychology to ascertain "how far different items of mental experience go together", and thereby to establish functional unities that are proven rather than assumed. He believes this need to have been successfully met by the device of correlational coefficients, and presents in his three final chapters an excellent résumé of the experiment and theory to which this search for functional unity has given rise. These chapters are respectively entitled 'Discovery of "G"', 'Specific Factors in Ability', and 'Orectic Factors'; taken together, they furnish what is certainly the best short account of recent discoveries in the field of factor-analysis. It is singularly appropriate that this account should have come from the pen of Prof. Spearman himself.

Psychology down the Ages is difficult to appraise as a whole. Can we fairly say that its author has fulfilled his ambitious mission of portraying 'the whole tide' of psychological development? Those who believe with Boring that the application to psychology of experimental method has been "the great outstanding event in the history of the study of mind" will assuredly be distressed by the relative unimportance of the role allotted to experiment in Prof. Spearman's chronicle. Others may feel aggrieved at the somewhat scanty treatment accorded to the numerous outstanding theoretical movements of the present day. Can we, in fact, be absolutely sure that Prof. Spearman is himself altogether free from confusion between the 'waves' and the 'tide' of psychological development? Yet taken all in all, *Psychology down the Ages* is a remarkable achievement. Prof. Spearman's prodigious erudition, the originality of his treatment, and the forcefulness of his style combine to produce one of the most stimulating and instructive psychological works of recent years. And whether the 'tide' of psychology be going out or coming in, Prof. Spearman will always be remembered as the creator of one of its largest and most influential 'waves'.

PUBLICATIONS RECENTLY RECEIVED

Visual Perception. By M. D. VERNON. Cambridge University Press. 1937. Pp. xi+247. 15s. net.

Unlike the *Gestalt* psychologists, Miss Vernon evidently does not believe the time to be ripe for the introduction of a deductive, still less of a 'hypothetico-deductive', methodology into the experimental study of perception. She proceeds throughout by induction: an enormous number of experimental results are systematically summarized and collated, and the validity of existing theory impartially assessed on the basis of its explanatory capacity.

Miss Vernon introduces her study with a very well-documented account of tachistoscopic perception under varying temporal conditions. Although believing perception to comprise an essentially unitary activity, she nevertheless inclines to the view that this activity may be resolved into a number of fundamental stages, each of which is to be envisaged as a constituent aspect of the total process. Such a view would seem to involve tacit acceptance of the assumption that a perceptual response elicited under 'normal' conditions of presentation is in some manner compounded out of such responses as may be shown to occur under exposure conditions of varying degree of brevity. Yet the author is evidently aware of this objection; she expressly stipulates that it would be unjustifiable to regard a perceptual process "thus analysed, interrupted or retarded" as necessarily identical with the process obtaining under rather more normal conditions.

Following a detailed and most serviceable account of experimental findings bearing upon the significance of attitudinal factors in perceiving, the author passes on to a discussion of some conditions determining phenomenal 'organisation'. In admirable chapters upon 'Figure' and 'Ground', Form Qualities, Apparent Movement and the Constancy Phenomena, she summarizes the principal *Gestalt* experiments, together with the results of researches rather more eclectic in character, whose implications are only too often overlooked by the *Gestalt* psychologists themselves. The author's general conclusion is that the experimental findings warrant at any rate partial acceptance of the configurationist position, but that integrated past experience must always be regarded as playing a very considerable part in the organization and differentiation of phenomenal data.

Perhaps the most useful (though also the most difficult) chapter is that which deals with the visual appreciation of spatial relations. Here again, the treatment may be considered by some to be too analytical, but no one could deny the immense service which the author has performed in systematizing the vast amount of hitherto diffuse experimental literature. This chapter should become the standard reference for those seeking an 'orientation' in this extraordinarily confusing and difficult field.

There is an interesting chapter upon the perceptual development of children, and Miss Vernon rounds off her study with an evaluation of the best-known typological approaches to the classification of perceptual and imaginal response. In this final chapter, the experimental observations relating to synaesthesia and kindred phenomena are briefly described, and an outline provided of Werner's theory of sensory differentiation.

The appendix on tachistoscropy is unique.

A Psychological Study of Religious Conversion. By W. LAWSON JONES. London: The Epworth Press. 1937. Pp. 397. 10s. 6d. net.

Although this book cannot be described as an easy one to read, partly because it is somewhat overloaded with references to a wide literature, and partly because

it often tends to mix up empirical remarks with sweeping generalizations, it is, emphatically, one that is worthy of serious study. Throughout the author is dealing with a mass of material, the great bulk of which he has collected directly for himself by means of a questionnaire, of a personal interview, or by information gathered from friends. His book is thus an original contribution to its subject, and although he shows himself to be almost devastatingly well acquainted with the works of other people, his conclusions are always derived primarily from his own empirically ascertained facts. A mere list of the main questions with which he deals cannot give any adequate idea of the book, but here they are: He lays down first his criteria for religious experience, and makes some clear remarks classifying the various general types of religious conversion. Then, turning to his actual data, and considering first the evidence in the case of 'gradual conversions', he presents his own facts with regard to the frequencies of such conversion for various age-groups, the characteristics of pre-conversion situations, the nature of conversion crises, and sums up the general balance of psychological factors involved. Next he discusses the ways in which conversion phenomena are expressed in 'behaviour, poise and attitude', both formally and in a more dynamic sense. There follows a series of studies of the same groups of questions with the reported cases of more sudden conversion. Throughout these and other parts of his work Mr Jones sets his own results into accurate and fruitful comparison with those of others. He now deals with a comparison of non-gradual conversions with relapsed cases, the latter having been characterized in the main by strong emotional crises only, without much re-direction of the cognitive and volitional life. Religious conversion is next compared with other forms of awakening, intellectual, aesthetic and the upspringing of love. A very interesting and important chapter deals with the social, physiological and temperamental settings of religious conversion. Psycho-analytic interpretations are examined with sympathy, but are not regarded as adequate to cover all the facts. The chapter in which conversion in Christian and in non-Christian communities is compared is of interest, but cannot be said to do more than introduce some wide and important problems.

The book as a whole is impressive. It is extremely well informed in general, and it is at the same time based firmly upon original research. Here and there the author shows a tendency to stray into more realms than seem necessary for the purposes of his argument or exposition. But when he does this it still remains that his remarks are usually interesting and provocative in themselves. In the actual writing, the book cannot challenge comparison with James's classic, but it contains a great deal more directly and honestly studied facts and it thoroughly deserves wide recognition as an excellent and reliable study in a difficult and complicated field.

Primitive Intelligence and Environment. By S. D. PORTEOUS. New York: The Macmillan Company. 1937. Pp. ix+325. 15s. net.

Although this is an interesting book to read, it does not greatly advance knowledge of the problems with which it is concerned. The first thirteen chapters are devoted to a consideration of the difficulties of comparative study in social psychology, with special reference to the use of tests, and to an account of the journeys undertaken and the peoples studied in the actual research. The remaining part of the book gives a rapid and, from a statistical point of view, insufficiently analysed presentation of the experiments and tests tried and of their results. Most of the discussion concerns some form or other of the well-known Porteous maze test, and comparative results are given—for the most part only in terms of average performance score—for a number of different 'racial' groups, especially for Australian Aborigines and certain Bantu people of South Africa. If this test really does possess some positive diagnostic value in respect to the capability of a native population to adapt itself to the demands of white culture, as the author tends to maintain, it is an extraordinary thing to find the Australian native coming out of it better than the Bantu. Other tests used were: psychophysical: brain capacity, strength of grip; measures of learning capacity:

Goddard form board and an assembling test; auditory and visual rote memory; the Porteous footprint matching technique and a short battery of intelligence tests of the normal kind. Actually very little can be concluded from the results as presented and it is not clear that any point of new or great methodological significance arises either.

Popular Psychological Fallacies. By JAMES G. TAYLOR. London: C. A. Watts and Co. 1938. Pp. vii + 275. 7s. 6d. net.

Here is a book that deserves a very cordial welcome and wide circulation. Mr Taylor is a Lecturer in Psychology at the University of Cape Town. What he has done is to take a number of statements about psychology or psychological problems which are given widespread expression and belief, and to show, always by reference to actual facts and empirical evidence, where and how they go astray. The result is one of those rather rare books which the psychologist who is well skilled in his subject can himself read with pleasure and profit, and then recommend with confidence to any intelligent reader who wants to find out something about psychology and its problems. A very wide range of problems is dealt with too. The first few have mainly to do with experimental topics, of perception, and practice, and thinking, and affection. Then Mr Taylor shows himself well equipped to consider applied topics, in industry, and in regard to mental health. He is not afraid to have a whole-hearted tilt at the psychoanalysts, and he goes to battle in a good-humoured as well as a sensible frame of mind. Finally he tackles some of the really difficult social problems of our time: racial psychology, dictatorships, the foundations of belief, and propaganda. Nobody is likely to agree with all of what he says, and that, of course, is a good thing; but nobody can fail to admire the courage and skill with which he goes to work. The book is very excellently and cheaply produced, and it should score a real success.

Coming into Being among the Australian Aborigines: An examination of all the evidence bearing upon the procreative beliefs of the Australian Aborigines. By M. F. ASHLEY-MONTAGU. London: George Routledge and Sons, Ltd. 1937. Pp. xxxv + 362. 21s. net.

It is possible that the long controversy which has been waged about the topic dealt with in this book will still continue. But it is very unlikely that anybody will in the future get a more complete collection of the relevant facts than has been made by Mr Ashley-Montagu. These facts are set forth clearly, convincingly, and with remarkably good arrangement, and the author is definitely sure that the Australian aboriginal has by himself no exact knowledge concerning the physiological relationship between intercourse and childbirth. What the nature of his views about both are is discussed at much length and with some reasonable speculation. The author is strongly opposed to the opinions of psycho-analytical anthropologists like Roheim. The book ought to become a classic in its own field, though whether the ultimate psychological significance of its problems merits so large and detailed a treatise is perhaps another matter. Prof. Malinowski contributes an appreciative preface.

Sex, Custom and Psychopathology: A Study of South African Pagan Natives. By B. J. F. LAUBSCHER. London: George Routledge and Sons, Ltd. 1937. Pp. xv + 347. 21s. net.

Dr Laubscher has made an extensive ethnological and psychiatric study of South Eastern Cape Bantu and in particular of the Tembu tribe. Nothing comparable either in thoroughness or authority has been issued hitherto, and the psychologist who is interested in many of the problems which are of vital importance in connexion with the effects of the impact of white civilization upon a relatively primitive people

will find much to enlighten and direct his studies. Rather more than the first half of the book is concerned more directly with anthropological than with psychological or psychiatric material. There are studies of folklore, of the practices and functions of native doctors, of customs connected with birth, childhood, adolescence, various rites and ceremonies, and particularly of initiation customs. Tembu marriages receive special attention, together with the institution of cousin marriage and practices of guardianship. Thereafter the psychological and especially the psychiatric interest of the work quickens. Dr Laubscher goes on to discuss the native's conception of mental disorder, sexual offences (studied mainly by means of a not altogether satisfactory questionnaire), and mental hospital material, especially as related to suicide, homicide and stock theft. The bulk of the natives actually dealt with by Dr Laubscher have been in more or less constant contact with whites. All the time one feels, and the author obviously feels also, the need for a still more thorough statistical and observational survey carried out by adequately trained psychologists and doctors. But the indications given, though they lack a very substantial basis, should serve to direct and to suggest further most important lines of investigation. "There is ample reason," says Dr Laubscher, "for thinking that psychopathological conditions are more prevalent among natives than among the Europeans." But he admits that this has not yet been proved, and if this important book stimulates the work that is urgently needed it will have served a valuable purpose.

A General Selection from the works of Sigmund Freud. Edited by JOHN RICKMAN. *Psycho-analytical Epitomes*, No. 1. London: The Hogarth Press and the Institute of Psycho-analysis. 1937. Pp. x+329. 5s. net.

Here is a book which can well be commended both to the student of psychology and to the intelligent general reader. Dr John Rickman has made an extremely judicious and admirable selection from Freud's own writings with the definite intent of displaying the "General Theory of Psycho-analysis". There are no case histories and, quite properly, there is no exposition of psycho-analytic technique. It seems perhaps rather a pity that some of the earlier ideas about mental structure and functions, and especially the views about memory which were fully expounded in *The Interpretation of Dreams*, could not find a place here. For they form a point of departure for the study of psycho-analysis as a theory which the ordinary student of psychology can well appreciate and is usually reasonably equipped to discuss. But as the editor points out, he has been able to include only about one-twentieth of Freud's work, and consequently cannot be severely criticized for what he has left out, though he is to be highly praised for what he has put in. It is a book which deserves a wide reading, and it should have a definite influence upon general understanding of Freud's constructive work, and upon the general attitude to his theories.

Love, Hate and Reparation. By MELANIE KLEIN and JOAN RIVIERE. *Psycho-analytical Epitomes*, No. 2. London: The Hogarth Press and the Institute of Psycho-analysis. 1937. Pp. vii+119. 3s. 6d. net.

"In this book", say the authors, "we discuss some aspects of the emotional life of men and women in civilized communities, the everyday manifestations of which are familiar to us all." An attempt is made to put this discussion into ordinary language, so that everybody can understand it. As the editor, Dr John Rickman, points out, this means that the detailed evidence for the views expounded must be omitted. He adds that this evidence "is alone convincing", but in this respect he is very likely wrong, for it means that the authors are able to state theoretical conclusions as if

they were wholly proved, and with an air of complete certainty which, as everybody knows, is very apt to produce acceptance. However this may be, the two lectures which are included, one on Hate, Greed and Aggression, by Joan Riviere, and the other on Love, Guilt and Reparation, by Melanie Klein, are both of them full of interest, both written with great lucidity, and both well fitted to appeal to the general reader. It is perhaps not unfair to say that, although the second lecture deals ostensibly with the constructive, co-operative and friendly impulses, there is throughout the whole volume rather a greater preoccupation with the destructive and negative forces in life, as if what are generally regarded as the happier sides of human existence are somewhat insecure. This may, of course, be true, or it may be a misinterpretation, since Miss Klein certainly says that "feelings of love and gratitude arise directly and spontaneously in the baby in response to the love and care of his mother". In any case, for the purpose for which they are meant, the lectures are undeniably good ones, and go further to indicate that this new series of psychological 'epitomes' will be one of much value.

Social Behaviour and Child Personality: An Exploratory Study of Some Roots of Sympathy. By LOIS BARCLAY MURPHY. New York: Columbia University Press; London: Humphrey Milford. 1937. Pp. ix+344. 17s. 6d. net.

This book describes a study of sympathetic behaviour in young children. The main part of the data was obtained from the continuous observation of two groups of children in nursery school playgrounds for daily two-hour periods over half a year. In addition there were obtained: information from the parents about the children's home behaviour; ratings from the teachers of various social behaviour traits; observations from 'framed', or experimentally controlled, situations suggesting sympathetic behaviour. It was found that some children showed high degrees of both aggressiveness and sympathy; others, a predominance of either aggressiveness or sympathy; and others, little of either and high degrees of fear and inhibition. But both the degree and type of sympathetic behaviour (active defence of other children, helping them, or comforting them actively or verbally) varied with the particular situation, its meaning to the child, the motivation involved and the social background, as well as with the child's total personality.

In fact, this study marks an interesting transition from the 'cross-section' method of studying the distribution of a given trait in a group (so popular in the U.S.A.) to the 'longitudinal section' method of studying its relation to and development within a total personality. The author seems reluctant to abandon the former method, though secretly convinced of the superiority of the latter. Had she been surer of her convictions, and concentrated on tracing out the general response tendencies of the individual manifested in long period fluctuations of behaviour, she might have avoided the slight confusion from which the book suffers; and have demonstrated more clearly her interesting conclusions as to the importance of 'range of behaviour'. It is to be hoped that she will develop this technique more fully in future studies.

The Parent-Child Relationship. By H. G. BAYNES, S. CROWN, S. H. LUBNER, A. C. COURT, M. MARCUS and F. G. CROOKSHANK. Individual Psychology Pamphlets, No. 17. London. 1937. Pp. 71. 2s. 6d.

This pamphlet contains reports of lectures on The Psychological Background of the Parent-Child Relationship, by Dr H. G. Baynes, and on Personality, Life-Style and the Demands of Life, by Dr F. G. Crookshank. There is, further, a report of

a symposium on Psychotherapy in General Practice by Drs S. Crown, S. H. Lubner, A. C. Court and M. Marcus.

Adult Attitudes to Children's Misdemeanours. By HELEN BOTT. University of Toronto Press. 1937. Pp. 21.

A number of common misdemeanours of children were studied by a method of paired comparison, and scale values were assigned them according to Thurstone's method of attitude measurement. Ten traits were assigned very nearly the same importance by groups of teachers, parents, public-health nurses, social workers and mental hygienists. These characters were: cruelty, deceit, sex troubles, stealing, bullying, sulkiness, tantrums, laziness, quarrelling and teasing, and they were placed in this order of seriousness. But while teachers considered disobedience, uncleanness, profanity and impertinence as of grave importance, mental hygienists thought these to be of small moment, and set above them avoidance of the group, daydreaming, shyness, lack of application, enuresis and mannerisms. Other, but smaller, disagreements appeared between the other group estimates. Whatever may be thought of the final forcefulness of the method, few people will be found to disagree with the general conclusion that the study points to "the need of agreement on fundamental issues among those responsible for the child's behaviour".

Emotional Episodes in the Child of School Age. By W. E. BLATZ, S. N. F. CHANT and M. D. SALTER. University of Toronto Press. 1937. Pp. 43.

This very interesting observational study is concerned mainly with emotional episodes in school life which have a maladaptive character. Study of the effects, frequency, exciting conditions and relation of these to such characters as chronological and mental maturity suggests strongly that what is most effective in overcoming the difficulties that this type of emotional episode raises is "facilitation of the learning of adaptive behaviour, intelligent guidance, reasonable routine and demands commensurate with maturity". Though nothing very new emerges, the monograph is a good statement of the problems involved and the methods of treatment of them that appear to be most securely based upon psychology.

An Analysis of the Social Contacts of Pre-school Children with the Aid of Motion Pictures. By K. S. BERNHARDT, DOROTHY A. MILLICHAMP, MARION W. CHARLES and MARY P. MCFARLAND. University of Toronto Press. 1937. Pp. 53.

The social contacts of fifteen children in the day nursery were observed in ten-minute records for each child once a month for four months. The analysis presented is based both on direct observation and on the treatment of motion pictures. The children were placed in three age groups: two-year olds, two-and-a-half-year olds and three-year olds. There appeared to be no important differences between the amount of social contacts spontaneously effected by these different age groups within the period during which they were studied. But the investigators maintain that their qualitative analysis reveals certain qualitative differences, mainly in a greater persistence of the kind of contact sought and in a greater complexity of social relationship as age slightly increases. Some attempt is also made to deal with group pattern and with the results of individual differences, but the main demand of the monograph is for more work.

The Development of Children's Concepts of Causal Relations. By JEAN MARQUIS DEUTSCHE. Institute of Child Welfare Monograph, No. 3. University of Minnesota Press; London: Humphrey Milford. 1937. Pp. x+104. 9s. net.

Seven hundred and thirty-two children between the ages of eight and sixteen answered two sets of questions involving causal relationships. The material thus obtained was analysed in many different ways, both quantitatively and qualitatively, and the general results are here fully set forth. Probably the one that is of most general interest is that Piaget's classification of seventeen types of causal thinking receives no confirmation. Dr Deutsche suggests four types, but without very much enthusiasm. There was no evidence that children's reasoning about causal relationships develops by stages, and it looks as if the way in which these relationships are regarded depends far more on tuition and experience than upon anything belonging to intelligence considered as an endowment or upon socio-economic status. A great many other conclusions are set forth, and the monograph is one of considerable interest; its English price, however, is excessive.

Children's Dreams. By C. W. KIMMINS. London: Geo. Allen and Unwin, Ltd. 1937. Pp. 121. 4s. 6d. net.

It was worth while publishing this book mainly for the actual record of dreams which it contains. A selection of dreams is recorded of children from five to fourteen years of age, of girls in secondary schools, of children in industrial and reformatory schools and of blind, deaf and physically defective children. Earlier chapters deal briefly, and in a somewhat conventional manner, with the nature of the dream, the daydream and the dreams of normal children. A final chapter comments very briefly on the educational value of the dream.

Growing Minds: An Introduction to Educational Psychology. By H. BOMPAS SMITH. University of London Press. 1937. Pp. viii+221. 5s. net.

This is a very well-informed book which anybody who is beginning to be interested in the applications of psychology to education could read both with pleasure and profit. It has two parts: the first deals with the growth of the boy's world, and the second with the growth of the boy's mind. While it can claim no particular originality, it is written in a pleasant manner, and with full references to the original works upon which its general conclusions and outlook are based.

The Golden Science Series, Book III. By E. V. M. KNIGHT. University of London Press. 1937. Pp. 264. 3s. net.

This is the third of a series of books intended to cover the needs of a science course for children of 11-14 years of age. It deals in an interesting manner with a great variety of different topics, and is abundantly and well illustrated. Chapters are included on the nervous system, on vision and hearing, and a few remarks are made about other human senses and their functions. The book is accurate and should be of wide use in schools. The author keeps the practical sides of her subject in mind all the time, and many interesting and simple experiments are suggested.

The Psychologist at Work. By M. R. HARROWER. London: Kegan Paul, Trench, Trubner and Co. Ltd. 1937. Pp. xiv + 184. 5s. net.

Dr Harrower has written her book "for the general public". At the same time there cannot be many people who, if they are not bothered by the somewhat intimate style—with masses of 'you's' and 'I's' and 'we's'—would not gain pleasure and profit from its perusal. The book is put forward as an introduction to experimental psychology, and it should excellently serve this purpose, for, as Prof. Koffka says in his appreciative preface, "it makes familiar events lose their character of being commonplace, and shows fascinating problems behind facts which are usually being taken for granted". Apart from brief general remarks, the actual topics dealt with are: vision in general, colour and sound experiences, certain relatively primitive forms of behaviour, the ways in which more complex behaviour develops, learning and remembering, and emotional aspects of life. A final chapter deals with experimental psychology in retrospect. The point of view adopted is very much in sympathy with *Gestalttheorie*.

Out of My Life and Work. By AUGUSTE FOREL. London: Geo. Allen and Unwin, Ltd. 1937. Pp. 352. 16s. net.

Forel's autobiography has been excellently translated by Bernard Miall. It is a full, perhaps almost an overfull, account of an interesting and influential life. Forel tells the story of his student years, with their developing interests in entomology and in medicine, with a little mild dissipation, and with a series of experiences during the Franco-Prussian War which set him strongly on the side of international peace for the rest of his life. He then relates the story of his prime, when his psychiatric interests prompted his studies of sex and of hypnotism, and when, together with his wife, he became a strong teetotaler and advocate of temperance reform. His later years were occupied with many far journeys and much social work in connexion with total abstinence, peace, the prevention of prostitution and other related problems. The Great War clouded his end, but neither it nor a disabling illness destroyed his belief in the high destiny and happy future of the human race.

Things I cannot Forget. By PHILIP BOSWOOD BALLARD. University of London Press. 1937. Pp. vi + 278. 8s. 6d. net.

Nobody who has met Dr Ballard and enjoyed the enlivening effect of his society and conversation will need any additional encouragement to buy and to read this book. Here, with unflinching good humour and any amount of wit, he tells the story of his life. He includes a number of short but vivid character sketches of many important and well-known people, especially of people who have played a part in educational developments in this country. The number of extremely good and amusing stories which are included is remarkable, and a great many of them are used to point some sensible and penetrating comment upon matters of psychological and educational interest. The book will in fact give enjoyment and profit to all who read it, and if it gets anything like its deserts they will be a very large number.

A Guide to Aesthetics. By ARAM TOROSSIAN. Stanford University Press; London: Humphrey Milford. 1937. Pp. 343. 17s. net.

The author of this book is a Professor of Architecture in the University of California. His volume is said to be "addressed primarily to beginners", but for this purpose it is overloaded and it will be a very patient beginner who reads it right through. On the other hand anybody, beginner or not, who dips into the book at almost any part of it will find much food for thought. Pretty well the whole field

of aesthetics, and all forms of aesthetic expression, are dealt with, and in most of the sections there is a fairly strong bias towards psychological treatment. The author tends to be dogmatic both as to facts and as to opinions, but gives abundant references where facts are concerned. There seems to be nothing remarkably original in the work, but it is a very good survey and written, for the most part, in an interesting manner. A selected bibliography is included. There are a few excellently reproduced illustrations.

Grundlinien der Kunstpsychologie. Von O. H. STERZINGER. Leipzig: Leykam-Verlag. 1938. S. xiv + 279. RM. 8.60.

This is the first of two projected volumes. It deals with the exterior content and form of artistic expression (*Die Sinneswelt*), and the second volume is to deal with the inner experiences involved. The first section, which discusses the claims of sensory stimuli in relation to art, takes in turn all the important sensorial fields, and indicates and illustrates how these claims are set up and perhaps realized in art. The whole discussion is interesting and original, for the attempt is consistently made to treat the sensory stimuli chiefly in relation to their characteristic groupings. The second main section deals with the effects of form and of the whole form considered as the basis of the artistic reaction. The general bent of this volume is psychological throughout, although the author points out that a psychology of art is not the same thing as a theory of aesthetic. Numerous pictures are provided to illustrate those parts of the argument which have to do with painting. There is a large bibliography, but the references are almost entirely to German publications. The volume is well worthy of the attention of any psychologist who is interested in the study of art.

L'Esprit et le Réel dans les Limites du Nombre et de la Grandeur. Par FRANCIS MAUGE. Paris: Félix Alcan. 1937. Pp. 367. 30 f.

L'Esprit et le Réel Perçu. Par FRANCIS MAUGE. Paris: Félix Alcan. 1937. Pp. 316. 30 f.

These two volumes together constitute a sustained and constructive attempt to develop the principles according to which the mind deals with external reality in its pursuit of truth. They are thus epistemological rather than psychological in aim, but anybody who finds rest and enjoyment in subtle and clever argument will discover much to attract him in these able books. The general problem is the age-old one of how it is that the mind comes to appreciate its external environment as possessing particular properties (intensity, duration, succession and so on) and to treat it as organized within certain categories. The general solution attempted is that this is not due to any universally imposed or respected convention, but is due to a constant struggle in which we have to engage in order to accommodate ourselves to the actual character of real objects without losing the integrity of our own mental processes. In the first volume the argument is worked out in respect to mathematical and modern physical modes of thinking, and in the second volume more direct problems of the assignment of sensory qualities and of the reference to the body are discussed. But no brief notice can do justice to M. Mauge's work which is interesting always and frequently brilliant.

La Part de l'Imagination. Par J. J. VAN BIERVLIET. Paris: Félix Alcan. 1937. Pp. 203. 30 f.

There are four chapters in this book. They deal with the birth of images, the factors which determine what among a whole mass of apparently possible images are going to be utilized by a given person, the life of images, constructive imagination and the advantages and disadvantages of the imagination. On the whole the book is rather disappointing. The author has done little or no original research, and most

readers will probably find the volume mainly useful for the excellent collection of illustrations and stories which it contains. Much is said that is acute and true in a direct descriptive sense, but comparatively little use seems to be made of it all in regard to the light which it may perhaps throw upon psychological problems.

New Frontiers of the Mind. By J. B. RHINE. London: Faber and Faber, Ltd. 1938. Pp. 298. 7s. 6d. net.

This book gives a popular account of the work on extra-sensory perception carried out up to date by Dr Rhine and his collaborators at Duke University. There seems to be little doubt that certain individuals, under favourable conditions, possess the power of obtaining better than chance (often much better than chance) success in guessing cards from a pack of twenty-five—five each of five different symbols. This occurs even when neither experimenter nor subject knows what any of the cards are till after the whole pack has been gone through; and the results appear to be statistically significant. Similar results were obtained telepathically, with the experimenter and subject in different rooms, the subject guessing as the experimenter examined each card in turn. From the fact that the accuracy of guessing did not vary with the distance of the rooms apart, Rhine argues that extra-sensory perception cannot be produced by any known form of physical radiation.

There seems little doubt, then, that extra-sensory perception can take place; but we must await further demonstration of its phenomenological nature, through qualitative as well as quantitative experiments, before we can assign to it any greater significance than, say, to synaesthesia.

Preliminary Studies of a Vaudeville Telepathist. By S. G. SOAL. University of London Council for Psychical Investigation, Bulletin III. London. 1937. Pp. 96. 5s. net.

Josef Kraus is a vaudeville artist who has achieved a great reputation for stage performances revolving about the finding of objects hidden by members of the audience, card guessing, and various forms of character delineation. He submitted to a very extensive series of tests admirably designed and carried out by Mr S. G. Soal and some assistants. In general the results show clearly that Kraus relies for success in the object-finding experiments mainly upon a high sensitivity to sensory, especially motor, cues given to him unconsciously by his audience. As regards the card guessing: "It is shown definitely that the figure on the face of the card plays no part in the process of recognition", but that Kraus can usually identify a card which he has previously handled. When tactual and visual cues are ruled out, he fails with the cards, just as he does with the hidden objects. As Mr Soal says: "These experiments suggest a most serious source of error in a great many of the card-guessing tests described by Ina Jephson and Dr J. B. Rhine." All the controlled experiments on telepathy, clairvoyance and character delineation produced negative results. The whole report forms an admirable illustration of experimental method applied in this field.

The Last Crossing. By GLADYS OSBORNE LEONARD. London: Cassell and Co., Ltd. 1937. Pp. 218. 7s. 6d. net.

Mrs Leonard is a well-known medium. This book is written about 'etheric bodies' and the like, and is based upon personal experiences.

Felicità. By R. E. SPENCER. London: Cassell and Co., Ltd. 1938. Pp. 157. 7s. 6d. net.

This is a novel about an old house, its former inhabitants, a sick man, and the ghost of a lady. It is all rather solemn.

Neur-Electro-Magnetic Theory of Nerve Reactions. By T. A. BENDRAT.
Beckley College, West Virginia. 1937. Pp. xxxiv + 122.

Dr Bendrat sets out to "explain the processes of nerve reactions as well as those involving thinking as purely physiological processes on the basis of electro-magnetic principles". It is clear that he has read widely in the recent work of electro-physiology. He also gives a large number of complicated and simple diagrams of a somewhat speculative character which are supposed to illustrate electrical processes in nerve units and groups. His contentions run far beyond anything that has as yet been proved.

The Theory and Practice of Yoga. By SARDAR S. SINGH. Czechoslovakia:
"Politica" Press. 1937. Pp. 134. 4s. net.

The Light of the Mind. By B. B. GATTELL and H. S. GATTELL. Phila-
delphia: Dorrance and Co. 1937. Pp. 140. \$1.75.

The Anatomy of Spirit: An Enquiry into the Origins of Religious Emotion.
By JACK LINDSAY. London: Methuen and Co., Ltd. 1937. Pp.
vii + 183. 5s. net.

Hume's Theory of Knowledge: A Critical Examination. By CONSTANCE
MAUND. London: Macmillan and Co. 1937. Pp. xxi + 310. 12s. 6d.
net.

Time, Cause and Eternity. The Forwood Lectures at Liverpool, 1935.
By J. L. STOCKS. With a foreword by the ARCHBISHOP OF YORK.
London: Macmillan and Co. 1938. Pp. 163. 6s. net.

The Concept of Morals. By W. T. STACE. London: Macmillan and Co.
1937. Pp. x + 307. 8s. 6d. net.

Toxicity of Industrial Solvents. Summaries of Published Work. By
ETHEL BROWNING. London: H.M. Stationery Office. Industrial
Health Research Board. 1937. Pp. 396. 7s. 6d. net.

